```
TCAAAACGTT AGAGAAAGTT CTTCAAAAGT GCAGCCCAGA GCTGCTGGGC CCACTGGCCG   3000
TCCTGCATTT CTGGTTCCCA GACCCCAATG CCTCCCATTC GGATGGATCT CTGCGTTTTT   3060
ATACTGAGTG TGCCTAGGTT GCCCCTTATT TTTTATTTTC CCTGTTGCGT TGCTATAGAT   3120
GAAGGGTGAG GACAATCGTG TATATGTACT AGAACTTTTT TATTAAAGAA A
```

A145 Protein sequence:
Gene name: Cadherin 3, P-cadherin (placental)
Unigene number: Hs.2877
Probeset Accession #:   X63629
Protein Accession #:   CAA45177
Signal sequence: 1-24
Transmembrane domain: 659-675

Cellular localization: plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
MGLPRGPLAS LLLLQVCWLQ CAASEPCRAV FREAEVTLEA GGAEQEPGQA LGKVFMGCPG   60
QEPALFSTDN DDFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG   120
KGPFPQRLNQ LKSNKDRDTK IFYSITGPGA DSPPEGVFAV EKSTGWLLLN KPLDREEIAK   180
YELFGHAVSE NGASVEDPMN ISIIVTDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT   240
DEDDAIYTYN GVVAYSIHSQ EPKDPHDLMF TIHRSTGTIS VISSGLDREK VPEYTLTIQA   300
TDMDGDGSTT TAVAVVEILD ANDNAPMFDP QKYEAHVPEN AVGHEVQRLT VTDLDAPNSP   360
AWRATYLIMG GDDGDHFTIT THPESNQGIL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL   420
PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGIPTGEPVC VYTAEDPDKE NQKISYRILR   480
DPAGWLAMDP DSGQVTAVGT LDREDEQFVR NNIYEVMVLA MDNGSPPTTG TGTLLLTLID   540
VNDHGPVPEP RQITICNQSP VRHVLNITDK DLSPHTSPFQ AQLTDDSDIY WTAEVNEEGD   600
TVVLSLKKFL KQDTYDVHLS LSDHGNKEQL TVIRATVCDC HGHVETCFGP WKGGFILPVL   660
GAVLALLFLL LVLLLVRKK RKIKEPLLLP EDDTRDNVFY YGEEGGGEED QDYDITQLHR   720
GLEARPEVVL RNDVAPTIIP TPMYRPRPAN PDEIGNFIIE NLKAANTDPT APPYDTLLVF   780
DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD
```

A146 DNA SEQUENCE:
Gene name: TTK protein kinase
Unigene number: Hs.169840
Probeset Accession #:   M86699
Nucleic Acid Accession #:   NM_003318
Coding sequence: 1026-3551 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
GGAATTCCTT TTTTTTTTTT TTTGAGATGG AGTTTCACTC TTGTTGGCCA GGCTGGAGTG   60
CAATGGCACA ATCTCAGCTT ACTGCAACCT CCGCCTCCCG GGTTCAAGCG ATTCTCCTGC   120
CTCAGCCTCC CAAGTAGCTG GGATTACAGG CATGTGCCAC CACCCCTGGC TAACTAATTT   180
CTTTTCTATT TAGTAGAGAT GGGGTTTCAC CATGTTGGTC AGGCTGGTCT TGAACTCCTG   240
ACCTCAGGTG ATCCACTTGC CTTGGCCTCC CAAAGTGCTA GGATTACAGC CGTGAAACTG   300
TGCCTGGCTG ATTCTTTTTT TGTTGTTGGA TTTTTGAAAC AGGGTCTCCC TTGGTCGCCC   360
AGGCTGGAGT GCAGTGGTGC GATCTTGGCT CACTATAACC TCCACCTCCT GGTTTCAAGT   420
GATCCTCCCA CTTTAGCCTC CTGAGTAGCT GTGATTACAG GCGTGCACCA CCACACCCGG   480
CTAATTTTTA TTATTTTTATT AGAGACAGGG TTTCACCATG TTGGCCAGGC TGTTCTCAAA   540
CTCCTGGACT CAAGGGATCC GCCTGCCTCC ACTTCCCAAA GTCCCGAGAT TACAGGTGTG   600
AGTCACCATG CCTGACCTTA TAATTCTTAA GTCATTTTTT CTGGTCCATT TCTTCCTTAG   660
GGTCCTCACA ACAAATCTGC ATTAGGCGGT ACAATAATCC TTAACTTCAT GATTCACAAA   720
AGGAAGATGA AGTGATTCAT GATTTAGAAA GGGGAAGTAG TAAGCCCACT GCACACTCCT   780
GGATGATGAT CCTAAATCCA GATACAGTAA AAATGGGGTA TGGGAAGGTA GAATACAAAA   840
TTTGGTTTAA ATTAATTATC TAAATATCTA AAAACATTTT TGGATACATT GTTGATGTGA   900
ATGTAAGACT GTACAGACTT CCTAGAAAAC AGTTTGGGTT CCATCTTTTC ATTTCCCCAG   960
TGCAGTTTTC TGTAGAAATG GAATCCGAGG ATTTAAGTGG CAGAGAATTG ACAATTGATT   1020
CCATAATGAA CAAAGTGAGA GACATTAAAA ATAAGTTTAA AAATGAAGAC CTTACTGATG   1080
AACTAAGCTT GAATAAAATT TCTGCTGATA CTACAGATAA CTCGGGAACT GTTAACCAAA   1140
TTATGATGAT GGCAAACAAC CCAGAGGACT GGTTGAGTTT GTTGCTCAAA CTAGAGAAAA   1200
ACAGTGTTCC GCTAAGTGAT GCTCTTTTAA ATAAATTGAT TGGTCGTTAC AGTCAAGCAA   1260
TTGAAGCGCT TCCCCCAGAT AAATATGGCC AAAATGAGAG TTTTGCTAGA ATTCAAGTGA   1320
GATTTGCTGA ATTAAAAGCT ATTCAAGAGC CAGATGATGC ACGTGACTAC TTTCAAATGG   1380
CCAGAGCAAA CTGCAAGAAA TTTGCTTTTG TTCATATATC TTTTGCACAA TTTGAACTGT   1440
CACAAGGTAA TGTCAAAAAA AGTAAACAAC TTCTTCAAAA AGCTGTAGAA CGTGGAGCAG   1500
TACCACTAGA AATGCTGGAA ATTGCCCTGC GGAATTTAAA CCTCCAAAAA AAGCAGCTGC   1560
TTTCAGAGGA GGAAAAGAAG AATTTATCAG CATCTACGGT ATTAACTGCC CAAGAATCAT   1620
TTTCCGGTTC ACTTGGGCAT TTACAGANTA GGAACAACAG TTGTGATTCC AGAGGACAGA   1680
CTACTAAAGC CAGGTTTTTA TATGGAGAGA ACATGCCACC ACAAGATGCA GAAATAGGTT   1740
ACCGGAATTC ATTGAGACAA ACTAACAAAA CTAAACAGTC ATGCCCATTT GGAAAGTCC   1800
CAGTTAACCT TCTAAATAGC CCAGATTGTG ATGTGAAGAC AGATGATTCA GTTGTACCTT   1860
GTTTTATGAA AAGACAAACC TCTAGATCAG AATGCCGAGA TTTGGTTGTG CCTGGATCTA   1920
AACCAAGTGG AAATGATTCC TGTGAATTAA GAAATTTAAA GTCTGTTCAA AATAGTCATT   1980
TCAAGGAACC TCTGGTGTCA GATGAAAAGA GTTCTGAACT TATTATTACT GATTCAATAA   2040
CCCTGAAGAA TAAAACGAGA TCAAGTCTTC TAGCTAAATT AAAGAGTATC   2100
AAGAACCAGA GGTTCCAGAG AGTAACCAGA AACAGTGGCA AGCTAAGAGA AGTCAGAGT   2160
GTATTAACCA GAATCCTGCT GCATCTTCAA ATCACTGGCA GATTCCGGAG TTAGCCCGAA   2220
AAGTTAATAC AGAGCAGAAA CATACCCACTT TTGAGCAACC TGTCTTTTCA GTTTCAAAAC   2280
AGTCACCACC AATATCAACA TCTAAATGGT TTGACCCAAA ATCTATTTGT AAGACACCAA   2340
```

```
      GCAGCAATAC CTTGGATGAT TACATGAGCT GTTTTAGAAC TCCAGTTGTA AAGAATGACT   2400
      TTCCACCTGC TTGTCAGTTG TCAACACCTT ATGGCCAACC TGCCTGTTTC CAGCAGCAAC   2460
      AGCATCAAAT ACTTGCCACT CCACTTCAAA ATTTACAGGT TTTAGCATCT TCTTCAGCAA   2520
      ATGAATGCAT TTCGGTTAAA GGAAGAATTT ATTCCATATT AAAGCAGATA GGAAGTGGAG   2580
      GTTCAAGCAA GGTATTTCAG GTGTTAAATG AAAAGAAACA GATATATGCT ATAAAATATG   2640
      TGAACTTAGA AGAAGCAGAT AACCAAACTC TTGATAGTTA CGGGAACGAA ATAGCTTATT   2700
      TGAATAAACT ACAACAACAC AGTGATAAGA TCATCCGACT TTATGATTAT GAAATCACGG   2760
      ACCAGTACAT CTACATGGTA ATGGAGTGTG GAAATATTGA TCTTAATAGT TGGCTTAAAA   2820
      AGAAAAAATC CATTGATCCA TGGGAACGCA AGAGTTACTG GAAAAAATATG TTAGAGGCAG   2880
      TTCACACAAT CCATCAACAT GGCATTGTTC ACAGTGATCT TAAACCAGCT AACTTTCTGA   2940
      TAGTTGATGG AATGCTAAAG CTAATTGATT TTGGGATTGC AAACCAAATG CAACCAGATA   3000
      CAACAAGTGT TGTTAAAGAT TCTCAGGTTG GCACAGTTAA TTATATGCCA CCAGAAGCAA   3060
      TCAAAGATAT GTCTTCCTCC AGAGAGAATG GGAAATCTAA GTCAAAGATA AGCCCCAAAA   3120
      GTGATGTTTG GTCCTTAGGA TGTATTTTGT ACTATATGAC TTACGGGAAA ACACCATTTC   3180
      AGCAGATAAT TAATCAGATT TCTAAATTAC ATGCCATAAT TGATCCTAAT CATGAAATTG   3240
      AATTTCCCGA TATTCCAGAG AAAGATCTTC AAGATGTGTT AAAGTGTTGT TTAAAAAGGG   3300
      ACCCAAAACA GAGGATATCC ATTCCTGAGC TCCTGGCTCA TCCATATGTT CAAATTCAAA   3360
      CTCATCCAGT TAACCAAATG GCCAAGGGAA CCACTGAAGA AATGAAATAT GTTCTGGGCC   3420
      AACTTGTTGG TCTGAATTCT CCTAACTCCA TTTTGAAAGC TGCTAAAACT TTATATGAAC   3480
      ACTATAGTGG TGGTGAAAGT CATAATTCTT CATCCTCCAA GACTTTTGAA AAAAAAGGG   3540
      GAAAAAAATG ATTTGCAGTT ATTCGTAATG TCAGATAGGA GGTATAAAAT ATATTGGACT   3600
      GTTATACTCT TGAATCCCTG TGGAAATCTA CATTTGAAGA CAACATCACT CTGAAGTGTT   3660
      ATCAGCAAAA AAAATTCAGT GAGATTATCT TTAAAAGAAA ACTGTAAAAA TAGCAACCAC   3720
      TTATGGCACT GTATATATTG TAGACTTGTT TTCTCTGTTT TATGCTCTTG TGTAATCTAC   3780
      TTGACATCAT TTTACTCTTG GAATAGTGGG TGGATAGCAA GTATATTCTA AAAAACTTTG   3840
      TAAATAAAGT TTTGTGGCTA AAATGA
```

Gene name: TTK protein kinase
Unigene number: Hs.169840
Probeset Accession #:   M86699
Protein Accession #:   NP_003309
Signal sequence: none found
Transmembrane domains: none found
Protein Kinase Domain: 510-775
Cellular Localization: cytoplasmic and nuclear

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      MNKVRDIKNK FKNEDLTDEL SLNKISADTT DNSGTVNQIM MMANNPEDWL SLLLKLEKNS   60
      VPLSDALLNK LIGRYSQAIE ALPPDKYGQN ESFARIQVRF AELKAIQEPD DARDYPQMAR  120
      ANCKKFAFVH ISFAQFELSQ GNVKKSKQLL QKAVERGAVP LEMLEIALRN LNLQKKQLLS  180
      EEEKKNLSAS TVLTAQESFS GSLGHLQNRN NSCDSRGQTT KARFLYGENM PPQDAEIGYR  240
      NSLRQTNRTK QSCPFGRVPV NLLNSPDCDV KTDDSVVPCF MKRQTSRSEC RDLVVPGSKP  300
      SGNDSCELRN LKSVQNSHFK EPLVSDEKSS ELIITDSITL KNKTESSLLA KLEETKEYQE  360
      PEVPESNQKQ WQAKRKSECI NQNPAASSNH WQIPELARKV NTEQKHTTFE QPVFSVSKQS  420
      PPISTSKWFD PKSICKTPSS NTLDDYMSCF RTPVVKNDPP PACQLSTFYG QPACFQQQQH  480
      QILATPLQNL QVLASSSANE CISVKGRIYS ILKQIGSGGS SKVFQVLNRK KQIYAIKYVN  540
      LEEADNQTLD SYRNEIAYLN KLQQHSDKII RLYDYEITDQ YIYMVMECGN IDLNSWLKKK  600
      KSIDPWERKS YWKNMLEAVH TIHQHGIVHS DLKPANFLIV DGMLKLIDFG IANQMQPDTT  660
      SVVKDSQVGT VNYMPPEAIK DMSSSRENGK SKSKISPKSD VWSLGCILYY MTYGKTPFQQ  720
      IINQISKLHA IIDPNHEIEF PDIPEKDLQD VLKCCLKRDP KQRISIPELL AHPYVQIQTH  780
      PVNQMAKGTT EEMKYVLGQL VGLNSPNSIL KAAKTLYEHY SGGESEMSSS SKTFEKKRGK  840
      K
```

Gene name: endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 7
Unigene number: Hs.258583
Probeset Accession #:   NM_012152
Nucleic Acid Accession #:   NM_012152
Coding sequence: 43-1104 (underlined sequences correspond to start and stop codons)

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      CTTCTTTAAA TTTCTTTCTA GGATGTTCAC TTCTTCTCCA CAATGAATGA GTGTCACTAT   60
      GACAAGCACA TGGACTTTTT TATAATAGG AGCAACACTG ATACTGTCGA TGACTGGACA  120
      GGAACAAAGC TTGTGATTGT TTTGTGTGTT GGGACGTTTT TCTGCCTGTT TATTTTTTTT  180
      TCTAATTCTC TGGTCATCGC GGCAGTGATC AAAAACAGAA AATTTCATTT CCCCTTCTAC  240
      TACCTGTTGG CTAATTTAGC TGCTGCCGAT TTCTTCGCTG GAATTGCCTA TGTATTCCTG  300
      ATGTTTAACA CAGGCCCAGT TTCAAAAACT TTGACTGTCA ACCGCTGGTT TCTCCGTCAG  360
      GGGCTTCTGG ACAGTAGCTT GACTGCTTCC CTCACCAACT TGCTGGTTAT CGCCGTGGAG  420
      AGGCACATGT CAATCATGAG GATGCGGGTC CATAGCAACC TGACCAAAAA GAGGGTGACA  480
      CTGCTCATTT TGCTTGTCTG GGCCATCGCC ATTTTTATGG GGGCGGTCCC CACACTGGGC  540
      TGGAATTGCC TCTGCAACAT CTCTGCCTGC TCTTCCCTGG CCCCCATTTA CAGCAGGAGT  600
      TACCTTGTTT TCTGGACAGT GTCCAACCTC ATGGCCTTCC TCATCATGGT TGTGGTGTAC  660
      CTGCGGATCT ACGTGTACGT CAAGAGGGAA ACCAACGTCT TGTCTCCGCA TACAAGTGGG  720
      TCCATCAGCC GCCGGAGGAC ACCCATGAAG CTAATGAAGA CGGTGATGAC TGTCTTAGGG  780
      GCGTTTGTGG TATGCTGGAC CCCGGGCCTG GTGGTTCTGC TCCTCGACGG CCTGAACTGC  840
      AGGCAGTGTG GCGTGCAGCA TGTGAAAAGG TGGTTCCTGC TGCTGGCGCT GCTCAACTCC  900
      GTCGTGAACC CCATCATCTA CTCCTACAAG GACGAGGACA TGTATGGCAC CATGAAGAAG  960
      ATGATCTGCT GCTTCTCTCA GGAGAACCCA GAGAGGCGTC CCTCTCGGCAT CCCCCTCCACA 1020
      GTCCTCAGCA GGAGTGACAC AGGCAGCCAG TACATAGAGG ATAGTATTAG CCAAGGTGCA 1080
```

```
GTCTGCAATA AAAGCACTTC CTAAACTCTG GATGCCTCTC GGCCCACCCA GGTGATGACT  1140
GTCTTAGG
```

A149 Protein sequence:

Gene name:	endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 7
Unigene number:	Hs.258583
Probeset Accession #:	NM_012152
Protein Accession #:	NP_036284
Signal sequence:	none found
Transmembrane domains:	31-53, 66-88, 150-172, 190-211, 239-261, 277-295
Cellular Localization:	plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
MNECHYDKHM DFFYNRSNTD TVDDWTGTKL VIVLCVGTFF CLFIFFSNSL VIAAVIKNRK  60
FHFPFYYLLA NLAAADFFAG IAYVFLMFNT GPVSKTLTVN RWFLRQGLLD SSLTASLTNL  120
LVIAVERHMS IMRMRVHSNL TKKRVILLIL LVWAIAIFMG AVPTLGNNCL CNISACSSLA  180
PIYSRSYLVF WTVSNLMAFL IMVVVYLRIY VYVKRKINVL SPHTSGSISR RRTFMKLMKT  240
VMTVLGAFVV CWTPGLVVLL LDGLNCRQCG VQHVKRWFLL LALLNSVVNP IIYSYKDEDM  300
YGTMKKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK STS
```

Prostate

A150 DNA SEQUENCE

Gene name:	ESTs
Unigene number:	Hs.293616
Probeset Accession #:	AW043782
Nucleic Acid Accession #:	none found

Coding sequence: 38-1075 (underlined sequence corresponds to start and stop codon)

```
1          11         21         31         41         51
|          |          |          |          |          |
AGCAACGACG CCGGGCAGCG GGAGCGGCGG CCGCGCCATG TGGCTGCTGG GGCCGCTGTG  60
CCTGCTGCTG AGCAGCGCCG CGGAGAGCCA GCTGCTCCCC GGGAACAACT TCACCAATGA  120
GTGCAACATA CCAGGCAACT TCATGTGCAG CAATGGACGG TGCATCCCGG CGCGCCTGGCA  180
GTGTGACGGG CTGCCTGACT GCTTCGACAA GAGTGATGAG AAGGAGTGCC CCAAGGCTAA  240
GTCGAAATGT GGCCCAACCT TCTTCCCCTG TGCCAGCGGC ATCCATTGCA TCATTGGTCG  300
CTTCCGGTGC AATGGGTTTG AGGACTGTCC CGATGGCAGC GATGAAGAGA ACTGCACAGC  360
AAACCCTCTG CTTTGCTCCA CCGCCCGCTA CCCACTGCAAG AACGGCCTCT GTATTGACAA  420
GAGCTTCATC TGCGATGGAC AGAATAACTG TCAAGACAAC AGTGATGAGG AAAGCGTGTGA  480
AAGTTCTCAA GAACCCGGCA GTGGGCAGGT GTTTGTGACT TCAGAGAACC AACTTGTGTA  540
TTACCCCAGC ATCACCTATG CCATCATCGG CAGCTCCGTC ATTTTTGTGC TGGTGGTGGC  600
CCTGCTGGCA CTGGTCTTGC ACCACCAGCG GAAGCGGAAC AACCTCATGA CGCTGCCCGT  660
GCACCGGCTG CAGCACCCTG TGCTGCTGTC CCGCCTGGTG GTCCTGGACC ACCCCCACCA  720
CTGCAACGTC ACCTACAACG TCAATAATGG CATCCAGTAT GTGGCCAGCC AGGCGGAGCA  780
GAATGCGTCG GAAGTAGGCT CCCCCACCCTC CTACTCCGAG GCCTTGCTGG ACCAGAGGCC  840
TGCGTGGTAT GACCTTCCTC CACCGCCCTA CTCTTCTGAC ACGGAATCTC TGAACCAAGC  900
CGACCTGCCC CCCTACCGCT CCCGGTCCGG GAGTGCCAAC AGTGCCAGCT CCCAGGCAGC  960
CAGCAGCCTC CTGAGCGTGG AAGACACCAG CCACAGCCCG GGGCAGCCTG GCCCCCAGGA  1020
GGGCACTGCT GAGCCCAGGG ACTCTGAGCC CAGCCAGGGC ACTGAAGAAG TATAAGTCCC  1080
AGTTATTCCA AAGTCCATAT GGGTTAATCT GCTCTGACTT GTTGCCATTC TAACAATTTG  1140
TGCTCATGGG AAGCTCTTTA AGCACCTGTA AGGATGTCTC AAGTTACAGT TTGGGATATT  1200
AACTATCTCT GCATTCCCCT CCTCCCCCAG ACTTCAGAGA TGTTTTTCTG GCGTCTCAGT  1260
TGACATGATC TGTTGTGCGT CTTTCTGTC AGGTCACTCT TCCCTTGGGA CCCGAGATCA  1320
CACCCTCATT TTTCACATTA TTCTGTTTCT GTTGGAGAGA CAGCATATAA AACAGTATTG  1380
AAATAGGCTG GGAGAGAGCA ATGTTTCTGT GCTATATTGG ATGCTGAGCA GTGCAGGAGA  1440
CGCTGGACCC AATTCTCTCT GCTGGGTAGT TACCTTATAG CATTTGGGGA TTTGGGTTAG  1500
ATGATCTAAC CAGGAGGCCA TCACTGGATG GTCACCCCCC CAAAAAAATT CCATTTGAGC  1560
ATCAAAACCT GCTTTGCACA ATCCTATTTG ATGCCCCCAG TTCAGCAGAG TCAGTGGCCA  1620
AAGAAAACTT TGGACGTGAG TAACACCCTT CAGCAGTCGC AACGTTATTT TGGTTTTGTG  1680
AAGGACTCTG AAACCATCTA CCCTGTATAA ATTCTGGCTT TAGAAATTTG CCCAAGAATG  1740
CTCATTCTGA GAGCTTTCCT CAGCAGCATA TATCATCAGC CTCATCCTAA AATAGGCAGG  1800
GAGCCCCTCC CATGAGTTTA TCCAAGTTCT CAGCTCCTAA AATGCAGGCT GCCAAGACCC  1860
TACACCTGCC CTGGCTCTAC AGCCACTTAC CTGGTTTCTG GACTGTCACC CTCCCAGCTG  1920
ACCTGCCCGT AGCCAAGGAA TGAGGACCTA ACTTGAGTTG GCCCAAAGTC TGACCTGGCT  1980
GTATGTCCCT GTGGCCCACA CCCAGCCTGT CTTGCTCATT CATGCAGCCT CAACACTGGC  2040
CTCCAAAGTT CCCTTAACAC TTGCAAAGTC CTTTTTACCT GTGCATTTGG ACTTGAGGAC  2100
ACTGGTTTCT ATCACAGGTG AGAGCCATGT TCAATACCTC CAGCAAGCTC TCCTGGCTCC  2160
CTGCACTGTG CACGCTCCTC TTCCCAAGGT CCCAATACCA GCACCTCTAG TTAGAGTTAG  2220
GGTCAGGGTC AGGCCTCTCC CAACATCCCA GTAGTTCTC CTCTGAGACA CATGGGCAAG  2280
AGACAATTTG GAGTCAAGAT TTTCCATTTG GATCTATTTT AAATCTTTTA GAAATGCATT  2340
TGAAACAGTG TGTTTGTTTT TTCCCTTCTA GTTAAGGGAC TATTTATATG TGTATAGGAA  2400
AGCTGTCTCT TTTTTTGTTT TTCCTTTAAC AAGGTCCAAA GAAAGATGCA AAAGGAGATC  2460
ACACCCTTGC CCCGCTGAGC CCCGTGATAA CAAGTCACTC CAGACTAACC TGTGTGCCAG  2520
ACATTTGTGC ATTGTTGCAC TTTGAGGTTA TTATTTATCA AGTTCTTGAA GGAAGCAGAA  2580
AGAGGGACTC CTCTCTCCCT CCGTGTATAG TCTCTATGTT TGTGCTAGTT TTTCTTTTTT  2640
TTCTCTGTGT CCAGTCAGCC ACAGGGCCCG CCTCCCTGCA GGAATAAGGG GTAAAACGTT  2700
AGGTGTTGTT TGGCAAGAAA CCACACTGAC TGATGAGGGG TAAAATGGAA CCAGGTAGAG  2760
CCACTCCGGG CAGCTGTCAC CCATTCAGAA CTTCTTTCCG CAGCCTGAAGA AATGTTCAGT  2820
```

```
AACCTGTTTG ACGCTAATTA AAACAGAGCC TGCAGGAAGT GGGGCTAAAG TGGCATTCAG   2880
TGATCCTGTT CTGTAGACTT TTCTTTCTTT TTTTAACCAA ATCCAAAGGA TGTTACAGAA   2940
AAGCTAGCCA CTGGTATTTT GTTTTGTTTA AAAAAAAAAA GAAAGAAAGA AAGAAAGAAA   3000
AACGGAAAGG AACCTAGCTG CCTGTATCTT TCATTTTTAA AATAGCACTT GAGTTATTTT   3060
CTGAGTAATC CAATAAAGAA CTTTTGATGA CAGCCAGAAT GTGTTAGAAC TCTGGCTGAA   3120
CATTTCATCT CCTGTGAGTC AGAAGGGCTT TATTTCTCCC TTTGATGGGG CCCCTTCTTC   3180
TTTCTGGTGT TCTGGAAGTT GTTTAGAGGA AAGAATTCTA ATTTTAATTA ATTGCGCAGT   3240
GAGTTAATCT CACTCGCTTT TCTGCTTCCA GGCATCTTAG GAAAAACAAA TGGTTTTAGT   3300
AGATAAGGGA TGCCTACTAA TGCTTTTTTA AAACAAACAG GGACATTTTT ATTATAGATT   3360
TGATTTTTTT AATGAATGTT TTTAAAAATA TATAAATAGG ACACCAAAGC GGCAGGGTTT   3420
TTTTTGGGGG GAGGGGGTTT GTTTTCCAAC TCAAGATGGC ACATTAGTGG CCAGCAATAT   3480
TTTTTAACTC ATTCCAACCA GGAAGCTTTT TTATACATTG CCTAAATCTA CGCCAACCAG   3540
AAAATAGTCT CATCTCTTTT TTTCTCAAAT GAGATCCGTG TTTTATTTTA GCATTAAATT   3600
AGTTACACTG TGATGACTGG CCTATTACCT GACTCAGCTC CCTCTACCTT GAAATTGACA   3660
TTTTTAAAAA ATGCAACTAA GTGGTTAATA GTGTGTGACG CTCAAAGTTA ATGTAAACTG   3720
GAAAGGTTGT GTGTCGTTGC TTTTTGTGTT TTGGTTAGGC TTGGTTTTGT TTTTTAATTT   3780
TTATACTTTC TAATAAATTT GCAGTTTCAT TCTTTCTGTT TGTGCAAAWG GWMCTAMARM   3840
AAMMAAAAAC AWYWTTGGGG GGGCTTGGGC CTCGGAAAAA GTTTTTAACA CCACTTCGGG   3900
TGGGGCGGCG GGGCCCACGT AGGTACGGCG ACCACGCGGG CCCAAACGGG ACCCCAGAAG   3960
GAAACCCTGG CCAAGAAAAA GGTGGCGAGA ATTCTCCACA CCAGAAAACA ACGCGCCGGG   4020
GGAAACCGCA GAGTGTTGCG TAAACCACAC CCGAAGAGAG AACTCAGAAG CACACAAGCG   4080
GGACTCAACC AGGAGGACCC AAGGGAACCC GATAGAGTAC G
```

A151 Protein sequence:

Gene name:	ESTs
Unigene number:	Hs.293616
Probeset Accession #:	AW043782
Protein Accession #:	none found
Signal sequence:	1-17
Transmembrane domains:	169-191
LDLa domains:	28-66, 70-108, 112-149
Cellular Localization:	plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
MWLLGPLCLL LSSAAESQLL PGNNFTNECN IPGNFMCSNG RCIPGAWQCD GLPDCFDKSD   60
EKECPKAKSK CGPTFFPCAS GIHCIIGRFR CNGFEDCPDG SDEENCTANP LLCSTARYEC  120
KNGLCIDKSF ICDGQNNCQD NSDEESCESS QEPGSGQVFV TSENQLVYYP SITYAIIGSS  180
VIFVLVVALL ALVLHHQRKR NNLMTLPVER LQHPVLLSRL VVLDHPHHCN VTYNVNNGIQ  240
YVASQABQNA SEVGSPPSYS EALLDQRPAW YDLPPPPYSS DTESLNQADL PPYRSRSGSA  300
NSASSQAASS LLSVEDTSHS PGQPGPQEGT AEPRDSEPSQ GTEEV
```

A152 DNA SEQUENCE

Gene name:	Transmembrane protease, serine 2 (TMPRSS2)
Unigene number:	Hs.105807
Probeset Accession #:	T48536
Nucleic Acid Accession #:	NM_005656.1

Coding sequence: 57-1535 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
GTCATATTGA ACATTCCAGA TACCTATCAT TACTCGATGC TGTTGATAAC AGCAAGATGG   60
CTTTGAACTC AGGGTCACCA CCAGCTATTG GACCTTACTA TGAAAACCAT GGATACCAAC  120
CGGAAAACCC CTATCCCGCA CAGCCCACTG TGGTCCCCAC TGTCTACGAG GTGCATCCGG  180
CTCAGTACTA CCCGTCCCCC GTGCCCCAGT ACGCCCCGAG GGTCCTGACG CAGGCTTCCA  240
ACCCCGTCGT CTGCACGCAG CCCAAATCCC CATCCGGGAC AGTGTGCACC TCAAAGACTA  300
AGAAAGCACT GTGCATCACC TTGACCCTGG GGACCTTCCT CGTGGGAGCT GCGCTGGCCG  360
CTGCCTACT CTGGAAGTTC ATGGGCAGCA AGTGCTCCAA CTCTGGGATA GAGTGCGACT  420
CCTCAGGTAC CTGCATCAAC CCCTCTAACT GGTGTGATGG CGTGTCACAC TGCCCCGGCG  480
GGGAGGACGA GAATCGGTGT GTTCGCTGTG ACGGACCAAA CTTCATCCTT CAGATGTACT  540
CATCTCAGAG GAAGTCCTGG CACCCTGTGT GCCAAGACGA CTGGAACGAG AACTACGGGC  600
GGGCGGCCTG CAGGGACATG GGCTATAAGA ATAATTTTTA CTCTAGCCAA GGAATAGTGG  660
ATGACAGCGG ATCCACCAGC TTTATGAAAC TGAACACAAG TGCCGGCAAT GTCGATATCT  720
ATAAAAAACT GTACCACAGT GATGCCTGTT CTTCAAAAGC AGTGGTTTCT TTACGCTGTT  780
TAGCCTGCGG GGTCAACTTG AACTCAAGGC GCCAGAGCAG GATCGTGGGC GGTGAGAGCG  840
CGCTCCCGGG GGCCTGGCCC TGGCAGGTCA GCCTGCACGT CCAGAACGTC CACGTGTGCG  900
GAGGCTCCAT CATCACCCCC GAGTGGATCG TGACAGCCGC CCACTGCGTG GAAAAACCTC  960
TTAACAATCC ATGGCATTGG ACGGCATTTG CGGGGATTTT GAGACAATCT TTCATGTTCT 1020
ATGGAGCCGG ATACCAAGTA CAAAAGTGA TTTCTCATCC AAATTATGAC TCCAAGACCA 1080
AGAACAATGA CATTGCGCTC ATGAAGCTCC AGAAGCCTCT GACTTTCAAC GACCTAGTGA 1140
AACCAGTGTG TCTGCCCAAC CCAGGCATGA TGCTGCAGCC AGAACAGCTC TGCTGGATTT 1200
CCGGGTGGGG GGCCACCGAG GAGAAAGGGA AGACCTCAGA AGTGCTGAAC GCTGCCAAGG 1260
TGCTTCTCAT TGAGACACAG AGATGCAACA GCAGATATGT CTATGACAAC CTGATCACAC 1320
CAGCCATGAT CTGTGCCGGC TTCCTGCAGG GGAACGTCGA TTCTTGCCAG GGTGACAGTG 1380
GAGGGCCTCT GGTCACTTCG AACAACAATA TCTGGTGGCT GATAGGGGAT ACAAGCTGGG 1440
GTTCTGGCTG TGCCAAAGCT TACAGACCAG GAGTGTACGG GAATGTGATG GTATTCACGG 1500
ACTGGATTTA TCGACAAATG AAGGCAAACG GCTAATCCAC ATGGTCTTCG TCCTTGACGT 1560
CGTTTTACAA GAAAACAATG GGGCTGGTTT TGCTTCCCCG TGCATGATTT ACTCTTAGAG 1620
ATGATTCAGA GGTCACTTCA TTTTTATTAA ACAGTGAACT TGTCTGGCTT TGGCACTCTC 1680
TGCCATACTG TGCAGGCTGC AGTGGCTCCC CTGCCCAGCC TGCTCTCCCT AAACCCCTTGT 1740
```

```
    CCGCAAGGGG TGATGGCCGG CTGGTTGTGG GCACTGGCGG TCAATTGTGG AAGGAAGAGG 1800
    GTTGGAGGCT GCCCCCATTG AGATCTTCCT GCTGAGTCCT TTCCAGGGGC CAATTTTGGA 1860
    TGAGCATGGA GCTGTCACTT CTCAGCTGCT GGATGACTTG AGATGAAAAA GGAGAGACAT 1920
5   GGAAAGGGAG ACAGCCAGGT GGCACCTGCA GCGGCTACCC TCTGGGGCCA CTTGGTAGTG 1980
    TCCCCAGCCT ACTTCACAAG GGGATTTTGC TGATGGGTTC TTAGAGCCTT AGCAGCCCTG 2040
    GATGGTGGCC AGAAATAAAG GGACCAGCCC TTCATGGGTG GTGACGTGGT AGTCACTTGT 2100
    AAGGGGAACA GAAACATTTT TGTTCTTATG GGGTGAGAAT ATAGACAGTG CCCTTGGTGC 2160
    GAGGGAAGCA ATTGAAAAGG AACTTGCCCT GAGCACTCCT GGTGCAGGTC TCCACCTGCA 2220
    CATTGGGTGG GGCTCCTGGG AGGGAGACTC AGCCTTCCTC CTCATCCTCC CTGACCCTGC 2280
10  TCCTAGCACC CTGGAGAGTG AATGCCCCTT GGTCCCTGGC AGGGCGCCAA GTTTGGCACC 2340
    ATGTCGGCCT CTTCAGGCCT GATAGTCATT GGAAATTGAG GTCCATGGGG GAAATCAAGG 2400
    ATGCTCAGTT TAAGGTACAC TGTTTCCATG TTATGTTTCT ACACATTGAT GGTGGTGACC 2460
    CTGAGTTCAA AGCCATCTT
```

15

25
```
    1          11         21         31         41         51
    |          |          |          |          |          |
    MALNSGSPPA IGFYYENHGY QPENPYPAQP TVVPTVYEVH PAQYYPSPVP QYAPRVLTQA  60
    SNPVVCTQPK SPSGTVCTSK TKKALCITLT LGTFLVGAAL AAGLLWKFMG SKCSNSGIEC 120
    DSSGTCINPS NWCDGVSHCP GGEDENRCVR LYGPNFILQM YSSQRKSWHP VCQDDWNENY 180
    GRAACRDMGY KNNFYSSQGI VDDSGSTSFM KLNTSAGNVD IYKKLYHSDA CSSKAVVSLR 240
    CLACGVNLNS SRQSRIVGGE SALPGAWPWQ VSLHVQNVHV CGGSIITPEW IVTAAHCVEK 300
    PLNNPWHWTA FAGILRQSFM FYGAGYQVQK VISHPNYDSK TKNNDIALMK LQKPLTFNDL 360
    VKPVCLPNPG MMLQPEQLCW ISGWGATEEK GKTSEVLNAA KVLLIETQRC NSRYVYDNLI 420
    TPAMICAGFL QGNVDSCQGD SGGPLVTSNN NIWWLIGDTS WGSGCAKAYR PGVYGNVMVF 480
    TDWIYRQMRA NG
```

```
    1          11         21         31         41         51
    |          |          |          |          |          |
    CAGAGAGGCT GTATTTCAGT GCAGCCTGCC AGACCTCTTC TGGAGGAAGA CTGGACAAAG   60
    GGGGTCACAC ATTCCTTCCA TACGGTTGAG CCTCTACCTG CCTGGTGCTG GTCACAGTTC  120
    AGCTTCTTCA TGATGGTGGA TCCCAATGGC AATGAATCCA GTGCTACATA CTTCATCCTA  180
    ATAGGCCTCC CTGGTTTAGA AGAGGCTCAG TTCTGGTTGG CCTTCCCATT GTGCTCCCTC  240
    TACCTTATTG CTGTGCTAGG TAACTTGACA ATCATCTACA TTGTGCGGAC TGAGCACAGC  300
    CTGCATGAGC CCATGTATAT ATTTCTTTGC ATGCTTTCAG GCATTGACAT CCTCATCTCC  360
    ACCTCATCCA TGCCCAAAAT GCTGGCCATC TTCTGGTTCA ATTCCACTAC CATCCAGTTT  420
    GATGCTTGTC TGCTACAGAT GTTTGCCATC CACTCCTTAT CTGGCATGGA ATCCACAGTG  480
    CTGCTGGCCA TGGCTTTTGA CCGCTATGTG GCCATCTGTC ACCCACTGCG CCATGCCACA  540
    GTACTTACGT TGCCTCGTGT CACCAAAATT GGTGTGGCTG CTGTGGTGCG GGGGGGCTGCA  600
    CTGATGGCAC CCCTTCCTGT CTTCATCAAG CAGCTGCCCT TCTGCCGCTC CAATATCCTT  660
    TCCCATTCCT ACTGCCTACA CCAAGATGTC ATGAAGCTGG CCTGTGATGA TATCCGGGTC  720
    AATGTCGTCT ATGGCCTTAT CGTCATCATC TCCGCCATTG GCCTGGCACTC ACTTCTCATC  780
    TCCTTCTCAT ATCTGCTTAT TCTTAAGACT GTGTTGGGCT TGACACGTGA AGCCCAGGCC  840
    AAGGCATTTG GCACTTGCGT CTCTCATGTG TGTGCTGTGT TCATATTCTA TGTACCTTTC  900
    ATTGGATTGT CCATGGTGCA TCGCTTTAGC AAGCGGCGTG ACTCTCCACT GCCCGTCATC  960
    TTGGCCAATA TCTATCTGCT GGTTCCTCCT GTGCTCAACC CAATTGTCTA TGGAGTGAAG 1020
    ACAAAGGAGA TTCGACAGCG CATCCTTCGA CTTTTCCATG TGGCCACACA CGCTTCAGAG 1080
    CCCTAGGGTGT CAGTGATCAA ACTTCTTTTC CATTCAGAGT CCTCTGATTC AGATTTTAAT 1140
    GTTAACATTT TGGAAGACAG TATTCAGAAA AAAAATTTCC TTAATAAAAA TACAACTCAG 1200
    ATCCTTCAAA TATGAAACTG GTTGGGGAAT CTCCATTTTT TCAATATTAT TTTCTTCTTT 1260
    GTTTTCTTGC TACATATAAT TATTAATACC CTGACTAGGT TGTGGTTGGA GGGTTATTAC 1320
    TTTTCATTTT ACCATGCAGT CCAAATCTAA ACTGCTTCTA CTGATGGTTT ACAGCATTCT 1380
    GAGATAAGAA TGGTACATCT AGAGAACATT TGCCAAAGGC CTAAGCACAG CAAAGGAAAA 1440
    TAAACACAGA ATATAATAAA ATGAGATAAT CTAGCTAAA CCTCTTCAGA 1500
    ACTCCCAACC ACATTGGATC TCAGAAAAAT ACTGTCTTCA AAATGACTTC TACAGAGAAG 1560
    AAATAATTTT TCCTCTGGAC ACTAGCACTT AAGGGGAAGA TTGGAAGTAA AGCCTTGAAA 1620
    AGAGTACATT TACCTACGTT AATGAAAGTT GACACACTGT TCTGAGAGTT TTCACAGCAT 1680
    ATGGACCCTG TTTTTCCTAT TTAATTTTCT TATCAACCCT TTAATTAGGC AAAGATATTA 1740
    TTAGTACCCT CATTGTAGCC ATGGGAAAAT TGATGTTCAG TGGGGATCAG TGAATTAAAT 1800
    GGGGTCATAC AAGTATAAAA ATTAAAAAAA AAAGACTTCA TGCCCAATCT CATATGATGT 1860
    GGAAGAACTG TTAAAGAGAC CAACAGGGTA GTGGGTTAGA GATTTCCAGA GTCTTACATT 1920
    TTCTAAAGGA GGTATTTAAT TTCTTCTCAC TCATCCAGTG TTGTATTTAG GAATTTCCTG 1980
    GCAACAGAAC TCATGGCTTT AATCCCACTA GCTATTGCTT ATTGTCCTGG TCCAATTGCC 2040
    AATTACCTGT GTCTTGGAAG AAGTGATTTC TAGGTTCACC ATTATGAAG ATTCTTATTC 2100
    AGAAAGTCTG CATAGGGCTT ATAGCAAGTT ATTTATTTTT AAAAGTTCCA TAGGTGTTTC 2160
    TGATAGGCAG TGAGGTTAGG GAGCCACCAG TTATGATGGG AAGTATGGAA TGGCAGGTGT 2220
    TGAAGATAAC ATTGGCCTTT TGAGTGTGAC TCGTAGCTGG AAAGTGAGGG AATCTTCAGG 2280
```

```
ACCATGCTTT ATTTGGGGCT TTGTGCAGTA TGGAACAGGG ACTTTGAGAC CGGGAAAGCA  2340
ATCTGACTTA GGCATGGGAA TCAGGCATTT TTGCTTCTGA GGGGCTATTA CCAAGGGGTA  2400
ATAGGTTTCA TCTTCAACAG GATATGACAA CAGTCTTAAC CAAGAAACTC AAATTACATA  2460
TACTAAAACA TGTGATCATA TATGTGGTAA GTTTCATTTT CTTTTTCAAT CCTCAGGTTC  2520
CCTGATATGG ATTCCTATNA CATGCTTTCA TCCCCTTTTG TAATGGATAT CATATTTGGA  2580
AATGCCTATT TAATACTTGT ATTTGCTGCT GGACTGTAAG CCCATGAGGG CACTGTTTAT  2640
TATTGAATGT CATCTCTGTT CATCATTGAC TGCTCTTTGC TCATCATTGA ATCCCCAGC   2700
AAAGTGCCTA GAACATAATA GTGCTTATGC TTGACACCGG TTATTTTTCA TCAAACCTGA  2760
TTCCTTCTGT GCTGAACACA TAGCCAGGCA ATTTTCCAGC CTTCTTTGAG TTGGGTATTA  2820
TTAAATTTTA GCCATTACTT CCAATGTGAG TGGAAGTGAC ATGTGCAATT TTTATACCTG  2880
GCTCATAAAA CCCTCCCATG TGCAGCCTTT CATGTTGACA TTAAATGTGA CTTGGGAAGC  2940
TATGTGTTAC ACAGAGTTAA TTAACCNGAA AGGCCTGGNA ATTTTTTGNN AANNAAACTG  3000
TGGCCNNGAG GCCCNCAACC CTTTTTNNNA ATTTGGCAAN NTCCCACTTT GTANTTTGGT  3060
AAGGAGGCCA GTTGGATAAG TGAAAAATAA AGTACTATTG TGTC
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MVDPNGNESS ATYFILIGLP GLEEAQFWLA FPLCSLYLIA VLGNLTIIYI VRTEESLHEP  60
MYIFLCMLSG IDILISTSSM PKMLAIFWFN STTIQFDACL LQMFAIHSLS GMESTVLLAM 120
AFDRYVAICH PLRHATVLTL PRVTKIGVAA VVRGAALMAP LPVFIKQLPF CRSNILSHSY 180
CLEQDVMKLA CDDIRVNVVY GLIVIISAIG LDSLLISFSY LLILKTVLGL TREAQAKAFG 240
TCVSHVCAVF IFYVPPIGLS MVHRFSKRRD SPLPVILANI YLLVPPVLNP IVYGVRTKEI 300
RQRILRLFHV ATHASEP
```

```
TCGGAGCCTG CGGAGGGTGG TGGTGGTGGT GGTGGTGGCC CTCGCCCGCC TCACTCATGC   60
CTCCTCCTCC TCTGCTCTCG CTCAGGCGCC TCGGTGGCGG TTGGTCGGCG GTTACGCGGC  120
TGGTGGTCGC GGCGGCCGGG GCTCGCTCTC GGGGAGGCCG GGGCGGATCT CGCGGCGCAG  180
GCGGCGGCTG CCGAGGTGGG GTCGCGCGGC GGAGGCGGCT CGAGCTTCGT GCTGCGCGCT  240
CGCTCTTGGG CTCCTCGCTG CAGGAGGAGT GTGACTATGT GCAGATGATC GAGGTGCAGC  300
ACAAGCAGTG CCTGGAGGAG GCCCAGCTGG AGAATGAGAC AATAGGCTGC AGCAAGATGT  360
GGGACAACCT CACCTGCTGG CCAGCCACCC CTCGGGGCCA GGTAGTTGTC TTGGGCCTGTC 420
CCCTCATCTT CAAGCTCTTC TCCTCCATTC AAGGCCGCAA TGTAAGCCGC AGCTGCACCG  480
ACGAAGGCTG GACGCACCTG GAGCCTGGCC CGTACCCCAT TGCCTGTGGT TTGGATGACA  540
AGGCAGCGAG TTTGATGAG CAGCAGACCA TGTTCTACGG TTCTGTGAAG ACCGGCTACA  600
CCATTGGCTA CGGCCTGTCC CTCGCCACCC TTCTGGTCGC CACAGCTATC CTGAGCCTGT  660
TCAGGAAGCT CCACTGTCAC CGGAACTACA TCCACATGCA CCTCTTCATA TCCTTCATCC  720
TGAGGGCTGC CGCTGTCTTC ATCAAAGACT TGGCCCTCTT CGACAGCGGG GAGTCGGACC  780
AGTGCTCCGA GGGCTCGGTG GGCTGTAAGG CAGCCATGGT CTTTTTCCAA TATTGTGTCA  840
TGGCTAACTT CTTCTGGCTG CTGGTGGAGG GCCTCTACCT GTACACCCTG CTTGCCGTCT  900
CCTTCTTCTC TGAGCGGAAG TACTTCTGGG GGTACATACT CATCGGCTGG GGGGTACCCA  960
GCACATTCAC CATGGTGTGG ACCATCGCCA GGATCCATTT TGAGGATTAT GGTCTGCTCA 1020
GGTGCTGGGA CACCATCAAC TCCTCACTGT GGTGGATCAT AAAGGGCCCC ATCCTCACCT 1080
CCATCTTGGT AAACTTCATC CTGTTTATTT GCATCATCCG AATCCTGCTT CAGAAACTGC 1140
GGCCCCAGA TATCAGGAAG AGTGACAGCA GTCCATACTC AAGGCTAGCC AGGTCCACAC 1200
TCCTGCTGAT CCCCCTGTTT GGAGTACACT ACATCATGTT CGCCTTCTTT CGGGACAATT 1260
TTAAGCCTGA AGTGAAGATG GTCTTTGAGC TCGTCGTGGG GTCTTTCCAG GGTTTTGTGG 1320
TGGCTATCCT CTACTGCTTC CTCAATGGTG AGGTGCAGGC GGAGCTGAGG CGGAAGTGGC 1380
GGCGCTGGCA CCTGCAGGGC GTCCTGGGCT GGAACCCCAA ATACCGGCAC CCGTCGGGAG 1440
GCAGCAACGG CGCCACGTGC AGCACGCAGG TTTCCATGGT GACCCGCGTC AGCCCAGGTG 1500
CCCGCCGCTC CTCCAGCTTC AAGCCGAAG TCTCCCTGGT CTGACCACCA GGATCCCAGC 1560
CCAAGCGGCC CCTCCCGCCC CTTCCCACTC GCAGCAGACG CCGGGGACAG AGGCCTGCCC 1620
GGGCGCGCCA GCCCCGGCCC TGGGCTCGGA GGCTGCCCCC CGCCCCCTGG TCTCTGGTCC 1680
GGACACTCCT AGAGAACGCA GCCCTAGAGC CTGCCTGGAG CGTTTCTAGC AAGTGAGAGA 1740
GATGGGAGCT CCTCTCCTGG AGGATGCAGG TGGAACTCAG TCATTAGACT CCTCCTCCAA 1800
AGGCCCCCTA CGCCAATCAA GGGCAAAAAG TCTACATACT TTCATCCTGA CTCTGCCCCC 1860
TGCTGGCTCT TCTGCCCAAT TGGAGGAAAG CAACCGGTGG ATCCTCAAAC AACACTGGTG 1920
TGACCTGAGG GCAGAAAGGT TCTGCCCGGG AAGGTCACCA GCACCAACAC CACGGTAGTG 1980
CCTGAAATTT CACCATTGCT GTCAAGTTCC TTTGGGTTAA GCATTACCAC TCAGGCATTT 2040
GACTGAAGAT GCAGCTCACT ACCCTATTCT CTCTTTACGC TTAGTTATCA GCTTTTTAAA 2100
GTGGGTTATT CTGGAGTTTT TGTTTGGAGA GCACACCTAT CTTAGTGGTT CCCCACCGAA 2160
GTGGACTGGC CCCTGGGTCA GTCTGGTGGG AGGACGGTGC AACCCAAGGA CTGAGGGACT 2220
CTGAAGCCTC TGGGAAATGA GAAGGCAGCC ACCAGCGAAT GCTAGGTCTC GGACTAAGCC 2280
TACCTGCTCT CCAAGTCTCA GTGGCTTCAT CTGTCAAGTG GGACTCTGTC ACACCAGCCA 2340
TTCTTATCTC TCTGTGCTGT GGAAGCAACA GGAATCAAGA GACTGCCCTC CTTGTCCACC 2400
CACCTATGTG CCAACTGTTG TAACTAGGCT CAGAGATGTG CACCCATGGG CTCTGACAGA 2460
```

```
AAGCAGATCC TCACCCTGCT ACACATACAG GATTTGAACT CAGATCTGTC TGATAGGAAT    2520
GTGAAAGCAC GGACTCTTAC TGCTAACTTT TGTGTATCGT AACCAGCCAG ATCCTCTTGG    2580
TTATTTGTTT ACCACTTGTA TTATTAATGC CATTATCCCT GAATTCCCCT TGCCACCCCA    2640
CCCTCCCTGG AGTGTGGCTG AGGAGGCCTC CATCTCATGT ATCATCTGGA TAGGAGCCTG    2700
CTGGTCACAG CCTCCTCTGT CTGCCCTTCA CCCCAGTGGC CACTCAGCTT CCTACCCACA    2760
CCTCTGCCAG AAGATCCCCT CAGGACTGCA ACAGGCTTGT GCAACAATAA ATGTTGGCTT    2820
GGAAAAAAAA AAAA
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MPPPPLLSLR RLGGGWSAVT RLVVAAAGAR SRGGRGGSRG AGGGGRGGVA RRRRLELRAA    60
RSLLGSSLQE ECDYVQMIEV QHKQCLEEAQ LENETIGCSK MWDNLTCWPA TPRGQVVVLA   120
CPLIFKLFSS IQGRNVSRSC TDEGWTHLEP GPYPIACGLD DKAASLDEQQ TMFYGSVKTG   180
YTIGYGLSLA TLLVATAILS LFRKLHCTRN YIHMHLFISF ILRAAAVFIK DLALFDSGES   240
DQCSEGSVGC KAAMVFFQYC VMANFFWLLV EGLYLYTLLA VSFFSERKYF WGYILIGWGV   300
PSTFTMVWTI ARIHFEDYGL LRCWDTINSS LWWIIKGPIL TSILVNFILF ICIIRILLQK   360
LRPPDIRKSD SSPYSRLARS TLLLIPLFGV HYIMFAFFPD NFKPEVKMVF ELVVGSFQGF   420
VVAILYCFLN GEVQAELRRK WRRWHLQGVL GWNPKYRHPS GGSNGATCST QVSMLTRVSP   480
GARRSSSFQA EVSLV
```

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAGCGGTG CGGGGGTGGC GGCTGGGACG CGGCCCCCCA GCTCGCCGAC CCCGGGCTCT     60
CGGCGCCGGC GCCAGCGCCC CTCTGTGGGC GTCCAGTCCT TGAGGCCGCA GAGCCCGCAG    120
CTCAGGCAGA GCGACCCGCA GAAACGGAAC CTGGACCTGG AGAAAAGCCT GCAGTTCCTG    180
CAGCAGCAGC ACTCGGAGAT GCTGGCCAAG CTCCATGAGG AGATCGAGCA TCTGAAGCGG    240
GAAAACAAGG GTGAGCCGGC GCGGGGCCCT AGGCCGGCCC TGCCTCCCCA GGCACACTCA    300
ACACTGCCGC TCCCGCAGCA CAGAAACACA GCCATCAACT CCAGCACACG CCTGGGCTCA    360
GGGGGAACAC AGGACGGGGA GCCCCTCCAG ACTGTCCTTG CCCACCTGGC TGCACTGGCC    420
CCTGTATGCC AACCCAGTGG GTACAGGTTC TGGGGGACCT GACAGATGC CGCTACCTCT     480
AGCCGTGGCT GGACGATGTT ATGCAGCCAA GCACAGCACG TGCTGCTCTC GGGAAGCCCA    540
GGGCCTGAGG TCATTGCAGG GCGGCAGGTG GCCACAGGGT GCTCCCCAGA CCTCCCTCCT    600
CCAAGTAGAG CTGAAATGGG AAGGAACCCC TGGGACAGCC CCTGCCCTGC TAGATCTTTG    660
CCTCAGATTG CTGCTGTGGC CAGGCCCAGG ATTTCCAGCC CTATGGCTCT GAGTCCTCAC    720
ATGCTGGGGG CCCAGGGGAT ATGGACACAC TCCATCCAGG GATCCCTTCC TGCCATCTGG    780
GCAGCAACCA TGGGGACAAA GGGAGGAAGC AGAGTCCTGT TTCCTTGCCA CTTGTCCAAG    840
GCACTTCCCC ATCCTGACAG CGGCCCCCAC CCAGCCCAGG ATCCTGGGCT GTGGTCTCAA    900
GCTCACTTCC CATTATCTTT GGGGCTGGGG CTGACATCAG GAGGACATCT GACTGGTGGA    960
TGGAGCCAGC CTGGGAACAT CGCCAGCTGG GCAGTGCCTA GGGCTCTCCC TTCCCAGGGA   1020
GACATGGAGA AGGGGGTTGA GGGAGGGGCC TTCCCTAGCC GCTGTGGCAA CTCCAGTGAG   1080
CTGTTCTGGG CAAAGTGTGG CCCAAGTCGG CAGCCCCAGC CCTGCAGTGC TGGGGACGCT   1140
GACAGGACAC GGGAAGAGGC CATGCTTTCC CTCGGGACCT GCTGTTCCAT GTGTCCCAAG   1200
CCCTCCTGCT TTCCAGATGG CCCCTCAGGA AACCACCTTT CTGCTCCCTTG TGCTCCCTTG   1260
GGCGCTCGCT GGGTCTGCAT CAACGGAGTG TGGGTAGAGC CGGGAGGACC CAGCCCTGCC   1320
AGGCTGAAGG AGGGCTCCTC ACGGACACAC AGGCCAGGAG GCAAGCGTGG GCGTCTTGCG   1380
GGCGGTAGCG CCGACACTGT GCGCTCTCCT GCAGACAGCC TCTCCATGTC AAGCTTCCAG   1440
TCTGTCAAGT CCATCTCTAA TTCAGCCAAC TCTCAAGGCA AGGCCAGGCC CCAGCCCGGC   1500
TCCTTCAACA AGCAAGATTC AAAAGCTGAC GTCTCCCAGA AGGCGGACCT GGAAGAGGAG   1560
CCCCTACTTC ACAACAGCAA GCTGGACAAA GTTCCTGGGG TACAAGGGCA GGCCAGAAAG   1620
GAGAAAGCAG AGGCCTCTAA TGCAGGAGCT GCCTGTATGG GGAACAGCCA GCACCAGGGC   1680
AGGCAGATGG GGGCGGGGGC ACACCCCCCA ATGATCCTGC CCCTTCCCCT GCGAAAGCCC   1740
ACCACACTTA GGCAGTGCGA AGTGCTCATC CGCGAGCTGT GGAATACCAA CCTCCTGCAG   1800
ACCCAAGAGC TGCGGCACCT CAAGTCCCTC CTGGAAGGGA GCCAGAGGCC CCAGGCAGCC   1860
CCGGAGGAAG CTAGCTTTCC CAGGGACCAA GAAGCCACGC ATTTCCCCAA GGTCTCCACC   1920
AAGAGCCTCT CCAAGAAATG CCTGAGCCCA CCTGTGGCGG AGCGTGCCAT CCTGCCCGCA   1980
CTGAAGCAGA CCCCGAAGAA CAACTTTGCC GAGAGGCAGA AGAGGCTGCA GGCAATGCAG   2040
AAACGGCGCC TGCATCGCTC AGTGCTTTGA
```

```
      1         11        21        31        41        51
      |         |         |         |         |         |
 5    MSGAGVAAGT RPPSSPTPGS RRRRQRPSVG VQSLRPQSPQ LRQSDPQKRN LDLEKSLQFL    60
      QQQHSEMLAK LHEEIEHLKR ENKGEPARGP RPALPPQAHS TLPLPQHRNT AINSSTRLGS   120
      GGTQDGEPLQ TVLAHLAALA PVCQPSGYRP WGTWTDAATS SRGWTMLCSQ AQHVLLSGSP   180
      GPEVIAGRQV ATGCSPDLPP PSRAEMGRNP WDSPCPARSL PQIAAVARPR ISSPMALSPH   240
      MLGAQGIWTH SIQGSLPAIW AATMGTKGGS RVLFPCHLSK ALPHPDSGPH PAQDPGLWSQ   300
10    AHPPLSLGLG LTSGGHLTGG WSQPGNIAAG AVPRALPSQG DMEKGVEGGP FPSRCGNSSE   360
      LPWAKCGPSR QPQPCSAGDA DRTREEAMLS LGTCCSMCPK PSCFPDGPSG NHLSRASAPL   420
      GARWVCINGV WVSPGGPSPA RLKEGSSRTH RPGGKRGRLA GGSADTVRSP ADSLSMSSFQ   480
      SVKSISNSAN SQGKARPQPG SPNKQDSKAD VSQKADLEEE PLLHNSKLDK VPGVQGQARK   540
15    EKAEASNAGA ACMGNSQHQG RQMGAGAHPP MILPLPLRKP TTLRQCEVLI RELWNTNLLQ   600
      TQELRHLKSL LEGSQRPQAA PEEASFPRDQ EATHPPKVST KSLSKKCLSP PVAERAILPA   660
      LKQTPKNNFA ERQKRLQAMQ KRRLHRSVL
```


A160 DNA SEQUENCE
Gene name: LIV-1 protein, estrogen regulated
Unigene number: Hs.79136
Probeset Accession #:   U41060
Nucleic Acid Accession #:   NM_012319.2
Coding sequence: 138-2405 (underlined sequences correspond to start and stop codons)

```
      1         11        21        31        41        51
      |         |         |         |         |         |
 30   CTCGTGCCGA ATTCGGCACG AGACCGCGTG TTCGCGCCTG GTAGAGATTT CTCGAAGACA    60
      CCAGTGGGCC CGTGTGGAAC CAAACCTGCG CGCGTGGCCG GGCCGTGGGA CAACGAGGCC   120
      GCGGAGACGA AGGCGCAATG GCGAGGAAGT TATCTGTAAT CTTGATCCTG ACCTTTGCCC   180
      TCTCTGTCAC AAATCCCCTT CATGAACTAA AAGCAGCTGC TTTCCCCCAG ACCACTGAGA   240
      AAATTAGTCA GAATTGGGAA TCTGGCATTA ATGTTGACTT GGCAATTTCC ACACGGCAAT   300
      ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAAATAA TTCTTTGTCA GTTGAAGGGT   360
      TCAGAAAATT ACTTCAAAAT ATAGGCATAG ATAAGATTAA AAGAATCCAT ATACACCATG   420
      ACCACGACCA TCACTCAGAC CACGAGCATC ACTCAGACCA TGAGCGTCAC TCAGACCATG   480
      AGCATCACTC AGACCACGAG CATCACTCTG ACCATGATCA TCACTCTCAC CATAATCATG   540
      CTGCTTCTGG TAAAAATAAG CGAAAAGCTC TTTGCCCAGA CCATGACTCA GATAGTTCAG   600
      GTAAAGATCC TAGAAACAGC CAGGGGAAAG GAGCTCACCG ACCAGAACAT GCCAGTGGTA   660
      GAAGGAATGT CAAGGACAGT GTTAGTGCTA GTGAAGCGAC CTCAACTGTG TACAACACTG   720
      TCTCTGAAGG AACTCACTTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAACTCTTCC   780
      CCAAAGATGT AAGCAGCTCC ACTCCACCCA GTGTCACATC AAAGAGCCGG GTGAGCCGGC   840
      TGGCTGGTAG GAAAACAAAT GAATCTGTGA GTGAGCCCCG AAAAAGGCTTT ATGTATTCCA   900
      GAAACACAAA TGAAAATCCT CAGGAGTGTT TCAATGCATC AAAGCTACTG ACATCTCATG   960
      GCATGGGCAT CCAGGTTCCG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA  1020
      TCAACCAAAT TGATGCTAGA TCTTGTCTGA TTCATACAAG TGAAAAGAAG GCTGAAATCC  1080
      CTCCAAAGAC CTATTCATTA CAAATAGCCT GGGTTGGTGG TTTTATAGCC ATTTCCATCA  1140
      TCAGTTTCCT GTCTCTGCTG GGGGTTATCT TAGTGCCTCT CATGAATCGG GTGTTTTTCA  1200
      AATTTCTCCT GAGTTTCCTT GTGGCACTGG CCGTTGGGAC TTTGAGTGGT GATGCTTTTT  1260
      TACACCTTCT TCCACATTCT CATGCAAGTC ACCACCATAG TCATAGCCAT GAAGAACCAG  1320
      CAATGGAAAT GAAAAGAGGA CCACTTTTCA GTCATCTGTC TTCTCAAAAC ATAGAAGAAA  1380
      GTGCCTATTT TGATTCCACG TGGAAGGGTC TAACAGCTCT AGGAGGCCTG TATTTCATGT  1440
      TTCTTGTTGA ACATGTCCTC ACATTGATCA AACAATTTAA AGATAAGGAG AAAAAGAATC  1500
      AGAAGAAACC TGAAAATGAT GATGATGTGG AGATTAAGAA GCAGTTGTCC AAGTATGAAT  1560
      CTCAACTTTC AACAAATGAG GAGAAAGTAG ATACGATGAA TCGAACTGAA GGCTATTTAC  1620
      GAGCAGACTC ACAAGAGCCC TCCCACTTTG ATTCTCAGCA GCCTGCAGTC TTGGAAGAAG  1680
      AAGAGGTCAT GATAGCTCAT GCTCATCCAC AGGAAGTCTA CAATGAATAT GTACCCAGAG  1740
      GGTGCAAGAA TAAATGCCAT TCACATTTCC ACGATACCAT CAGCCAGTCA GACGATCTCA  1800
      TTCACCACCA TCATGACTAC CATCATATTC TCCATCATCA CCACCACCAA AACCACCATC  1860
      CTCACAGTCA CAGCCAGCGC TACTCTCGGG AGGAGCTGAA AGATGCCGGC GTCGCCACTT  1920
      TGGCCTGGAT GGTGGATAATG GGTGATGGCC TGCACAATTT CAGCGATGGC CTAGCAATTG  1980
      GTGCTGCTTT TACTGAAGGC TTATCAAGTG GTTTAAGTAC TTCTGTTGCT GTGTTCTGTC  2040
      ATGAGTTGCC TCATGAATTA GGTGACTTTG CTGTTCTACT AAAGGCTGGC ATGACCGTTA  2100
      AGCAGGCTGT CCTTTATAAT GCATTGTCAG CCATGCTGGC GTATCTTGGA ATGGCAACAG  2160
      GAATTTTCAT TGGTCATTAT GCTGAAAATG TTTCTATGTG GATATTTGCA CTTACTGCTG  2220
      GCTTATTCAT GTATGTTGCT CTGGTTGATA TGGTACCTGA AATGCTGCAC AATGATGCTA  2280
      GTGACCATGG ATGTAGCCGC TGGGGGTATT TCTTTTTACA GAATGCTGGG ATGCTTTTGG  2340
      GTTTTGGAAT TATGTTACTT ATTTCCATAT TTGAACATAA AATCGTGTTT CGTATAAATT  2400
      TCTAGTTAAG GTTTAAATGC TAGAGTAGCT TAAAAGTTG TCATAGTTTC AGTAGGTCAT  2460
      AGGGAGATGA GTTTGTATGC TGTACTATGC AGCGTTTAAA GTTAGTGGGT TTTGTGATTT  2520
      TTGTATTGAA TATTGCTGTC TGTTACAAAG TCAGTTAAAG GTACGTTTTA ATATTTAAGT  2580
      TATTCTATCT TGGAGATAAA ATCTGTATGT GCAATTCACC GGTATTACCA GTTTATTATG  2640
      TAAACAAGAG ATTTGGCATG ACATGTTCTG TATGTTTCAG GGAAAAATGT CTTTAATGCT  2700
      TTTTCAAGAA CTAACACAGT TATTCCTATA CTGGATTTTA GGTCTCTGAA GAACTGCTGG  2760
      TGTTTAGGAA TAAGAATGTG CATGAAGCCT AAAATACCAA GAAAGCTTAT ACTGAATTTA  2820
      AGCAAAGAAA TAAAGGAGAA AAGAGAAGAA TCTGAGAATT GGGGAGGCAT AGATTCTTAT  2880
      AAAAATCACA AAATTTGTTG TAAATTAGAG GGGAGAAATT TAGAATTAAG TATAAAAAGG  2940
      CAGAATTAGT ATAGAGTACA TTCATTAAAC ATTTTTGTCA GGATTATTTC CCGTAAAAAC  3000
      GTAGTGAGCA CTCTCATATA CTAATTAGTG TACATTTAAC TTTGTATAAT ACAGAAATCT  3060
      AAATATATTT AATGAATTCA AGCAATATAC ACTTGACCAA GAAATTGGAA TTTCAAAATG  3120
      TTCGTGCGGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGTATCAC CAGACTGGGT  3180
      TATTGCCAAG TTATATATCA CCAAAAGCTG TATGACTGGA TGTTCTGGTT ACCTGGTTTA  3240
      CAAAATTATC AGAGTAGTAA AACTTTGATA TATATGAGGA TATTAAAACT ACACTAAGTA  3300
      TCATTTGATT CGATTCAGAA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCATTGT  3360
```

GAGCAATTGT CTTTATATAC GGTACTGTAG CCATACTAGG CCTGTCTGTG GCATTCTCTA   3420
GATGTTTCTT TTTTACACAA TAAATTCCTT ATATCAGCTT G

A161 PROTEIN SEQUENCE
Gene name: LIV-1 protein, estrogen regulated
Unigene number: Hs.79136
Protein Accession #:   NP_036451
Signal sequence: 1-21
Pfam domain: Zip[591-743]
Transmembrane domains: 330-346, 352-368, 427-444, 663-679, 688-703, 730-745
Cellular Localization: plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
MARKLSVILI LTFALSVTNP LHELKAAAFP QTTEKISPNW ESGINVDLAI STRQYHLQQL   60
FYRYGENNSL SVEGFRKLLQ NIGIDKIKRI HIHHDHDHHS DHEHHSDHER HSDHEHHSDH   120
EHHSDHDHHS HHNHAASGKN KRKALCPDHD SDSSGKDPRN SQGKGAHRPE HASGRRNVKD   180
SVSASEVTST VYNTVSEGTH FLETIETPRP GKLFPKDVSS STPPSVTSKS RVSRLAGRKT   240
NESVSEPRKG FMYSRNTMEN PQECFNASKL LTSHGMGIQV PLNATEFNYL CPAIINQIDA   300
RSCLIHTSEK KAEIPPKTYS LQIAHVGGFI AISIISFLSL LGVILVPLMN RVFFKPLLSF   360
LVALAVGTLS GDAFLHLLPH SHASHHHSHS HEEPAMEMKR GPLFSHLSSQ NIEESAYFDS   420
TWKGLTALGG LYFMFLVEHV LTLIKQFKDK KKKNQKKPEN DDDVEIKKQL SKYESQLSTN   480
EEKVDTDDRT EGYLRADSQE PSHFDSQQPA VLEEEEVMIA HAHPQEVYNE YVPRGCKNKC   540
HSHFHDTLGQ SDDLIHEHHD YHHILHHEHH QNHHPHSESQ RYSREELKDA GVATLAWMVI   600
MGDGLHNFSD GLAIGAAFTE GLSSGLSTSV AVFCHELPHE LGDFAVLLKA GMTVKQAVLY   660
NALSAMLAYL GMATGIFIGH YAENVSMWIF ALTAGLFMYV ALVDMVPEML HNDASDHGCS   720
RWGYFFLQNA GMLLGFGIML LISIFEHKIV FRINF
```

A162 DNA sequence
Gene name: bone morphogenetic protein receptor IB (ALK-6)
Unigene number: Hs.87223
Probeset Accession #:   AA250737
Nucleic Acid Accession #:   NM_001203
Coding sequence: 274-1782 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CGCGGGGCGC GGAGTCGGCG GGGCCTCGCG GGACGCGGGC AGTGCGGAGA CCGCGGCGCT   60
GAGGACGCGG GAGCCGGGAG CGCACGCGCG GGGTGGAGTT CAGCCTACTC TTTCTTAGAT   120
GTGAAAGGAA AGGAAGATCA TTTCATGCCT TGTTGATAAA GGTTCAGACT TCTGCTGATT   180
CATAACCATT TGGCTCTGAG CTATGACAAG AGAGGAAACA AAAGTTAAA CTTACAGCC    240
TGCCATAAGT GAGAAGCAAA CTTCCTTGAT AACATGCTTT TGCGAAGTGC AGGAAAATTA   300
AATGTGGGCA CCAAGAAAGA GGATGGTGAG AGTACAGCCC CCACCCCCCG TCCAAAGGTC   360
TTGCGTTGTA AATGCCACCA CCATTGTCCA GAAGACTCAG TCAACAATAT TTGCAGCACA   420
GACGGATATT GTTTCACGAT GATAGAAGAG GATGACTCTG GGTTGCCTGT GGTCACTTCT   480
GGTTGCCTAG GACTAGAAGG CTCAGATTTT CAGTGTCGGG ACACTCCCAT TCCTCATCAA   540
AGAAGATCAA TTGAATGCTG CACAGAAAGG AACGAATGTA ATAAAGACCT ACACCCTACA   600
CTGCCTCCAT TGAAAAACAG AGATTTTGTT GATGGACCTA TACACCACAG GGCTTTACTT   660
ATATCTGTGA CTGTCTGTAG TTTGCTCTTG GTCCTTATCA TATTATTTTG TTACTTCCGG   720
TATAAAGAC AAGAAACCGA ACCTCGATAC AGCATTGGGT TAGAACAGGA TGAAACTTAC   780
ATTCCTCCTG GAGAATCCCT GAGAGACTTA ATTGAGCAGT CTCAGAGCTC AGGAAGTGGA   840
TCAGGCCTCC CTCTGCTGGT CCAAAGGACT ATAGCTAAGC AGATTCAGAT GGTGAAACAG   900
ATTGGAAAAG GTCGCTATGG GGAAGTTTGG ATGGGAAAGT GGCGTGGCGA AAAGGTAGCT   960
GTGAAAGTGT TCTTCACCAC AGAGGAAGCC AGCTGGTTCA GAGAGACAGA AATATATCAG   1020
ACAGTGTTGA TGAGGCATGA AAACATTTTG GGTTTCATTG CTGCAGATAT CAAAGGGACA   1080
GGGTCCTGGA CCCAGTTGTA CCTAATCACA GACTATCATG AAAATGGTTC CCTTTATGAT   1140
TATCTGAAGT CCACCACCCT AGACGCTAAA TCAATGCTGA AGTTAGCCTA CTCTTCTGTC   1200
AGTGGCTTAT GTCATTTACA CACAGAAATC TTTAGTACTC AAGGCAAACC AGCAATTGCC   1260
CATCGAGATC TGAAAAGTAA AAACATTCTG GTGAAGAAAA ATGGAACTTG CTGTATTGCT   1320
GACCTGGGCC TGGCTGTTAA ATTTATTAGT GATACAAATG AAGTTGACAT ACCACCTAAC   1380
ACTCGAGTTG GCACCAAACG CTATATGCCT CCAGAAGTGT TGGACGAGAG CTTGAACAGA   1440
AATCACTTCC AGTCTTACAT CATGGCTGAC ATGTATAGTT TTGGCCTCAT CCTTTGGGAG   1500
GTTGCTAGGA GATGTGTATC AGGAGGTATA GTGGAAGAAT ACCAGCTTCC TTATCATGAT   1560
CTAGTGCCCA GTGACCCCTC TTATGAGGAC ATGAGGGAGA TTGTTGTGCAT CAAGAAGTTA   1620
CGCCCCTCAT TCCCAAACCG GTGGAGCAGT GATGAGTGTC TAAGGCAGAT GGGAAAACTC   1680
ATGACAGAAT GCTGGGCTCA CAATCCTGCA TCAAGGACAG CAGCCCTGCG GGTTAAGGAA   1740
ACACTTGCCA AAATGTCAGA GTCCCAGGAC ATTAAACTCT GATAGGGAGA GAAAAGTAAG   1800
CATCTCTGCA GAAAGCCAAC AGGTACTCTT CTGTTTGTGG GCAGAGCAAA AGACATCAAA   1860
TAAGCATCCA CAGTACAAGC CTTGAACATC GTCCTGCTTC CCAGTGGGTT CAGACCTCAC   1920
CTTTCAGGGA GCGACCTGGG CAAAGACAGA GAAGCTCCCA GAAGGAGAGA TTGATCCGTG   1980
TCTGTTTGTA GGCGGAGAAA CCGTTGGGTA ACTTGTTCAA GATATGATGC AT
```

A163 Protein sequence
Gene name: bone morphogenetic protein receptor IB (ALK-6)
Unigene number: Hs.72472 / Hs.87223
Probeset Accession #:   AA250737 / U89326
Protein Accession #:   NP_001194
Signal sequence: 1-13
Transmembrane domains: 128-144
PFAM domains: activin_receptor [30-111], protein kinase [204-491]
Cellular Localization: plasma membrane

```
     1          11         21         31         41         51
     |          |          |          |          |          |
  MLLRSAGKLN VGTKKEDGES TAPTPRPKVL RCKCHHCPE DSVNNICSTD GYCFTMIEED   60
  DSGLPVVTSG CLGLEGSDFQ CRDTPIPHQR RSIECCTERN ECNKDLHPTL PPLKNRDFVD  120
  GPIHHRALLI SVTVCSLLLV LIILFCYFRY KRQSTRPRYS IGLEQDETYI PPGESLRDLI  180
  EQSQSGSGS GLPLLVQRTI ARQIQMVKQI GKGRYGEVWM GKWRGEKVAV KVFFTEBAS   240
  WPRETEIYQT VLMRHENILG FIAADIKGTG SWTQLYLITD YHENGSLYDY LKSTTLDAKS  300
  MLKLAYSSVS GLCHLHTEIP STQGKPAIAH RDLKSKNILV KKNGTCCIAD LGLAVKFISD  360
  TNEVDIPPNT RVGTKRYMPP EVLDESLNRN HFQSYIMADM YSFGLILWEV ARRCVSGGIV  420
  EEYQLPYHDL VPSDPSYEDM REIVCIKKLR PSFPNRWSSD ECLRQMGKLM TECWAHNPAS  480
  RLTALRVKKT LAKMSESQDI KL
```

```
     1          11         21         31         41         51
     |          |          |          |          |          |
  GACAGTGTTC GCGGCTGCAC CGCTCGGAGG CTGGGTGACC CGCGTAGAAG TGAAGTACTT   60
  TTTTATTTGC AGACCTGGGC CGATGCCGCT TTAAAAAACG CGAGGGGCTC TATGCACCTC  120
  CCTGGCGGTA GTTCCTCCGA CCTCAGCCGG GTCGGGTCGT GCCGCCCTCT CCCAGGAGAG  180
  ACAAACAGGT GTCCCACGTG GCAGCCGCGC CCCGGGCGCC CCTCCTGTGA TCCCGTAGCG  240
  CCCCCTGGCC CGAGCCGCGC CCGGGTCTGT GAGTAGAGCC GCCCGGGCAC CGAGCGCTGG  300
  TCGCCGCTCT CCTTCCGTTA TATCAACATG CCCCCTTTCC TGTTGCTGGA GGCCGTCTGT  360
  GTTTTCCTGT TTTCCAGAGT GCCCCCATCT CTCCCTCTCC AGGAAGTCCA TGTAAGCAAA  420
  GAAACCATCG GGAAGATTTC AGCTGCCAGC AAAATGATGT GGTGCTCGGC TGCAGTGGAC  480
  ATCATGTTTC TGTTAGATGG GTCTAACAGC GTCGGGAAAG GGAGCTTTGA AAGGTCCAAG  540
  CACTTTGCCA TCACAGTCTG TGACGGTCTG GACATCAGCC CCGAGAGGGT CAGAGTGGGA  600
  GCATTCCAGT TCAGTTCCAC TCCTCATCTG GAATTCCATT TGGATTCATT TTCAACCCAA  660
  CAGGAAGTGA AGGCAAGAAT CAAGAGGATG GTTTTCAAAG GAGGGCGCAC GGAGACGGAA  720
  CTTGCTCTGA AATACCTTCT GCACAGAGGG TTGCCTGGAG GCAGAAATGC TTCTGTGCCC  780
  CAGATCCTCA TCATCGTCAC TGATGGGAAG TCCCAGGGGG ATGTGGCACT GCCATCCAAG  840
  CAGCTGAAGG AAAGGGGTGT CACTGTGTGTT GCTGTGGGGG TCAGGTTTCC CAGGTGGGAG  900
  GAGCTGCATG CACTGGCCAG CGAGCCTAGA GGGCAGCACG TGCTGTTGGC TGAGCAGGTG  960
  GAGGATGCCA CCAACGGCCT CTTCAGCACC CTCAGCAGCT CGGCCATCTG CTCCAGCGCC 1020
  ACGCCAGACT GCAGGGTCGA GGCTCACCCC TGTGAGCACA GGACGCTGGA GATGGTCGGG 1080
  GAGTTCGCTG GCAATGCCCC ATGCTGGAGA GGATCGCGGC GGACCCTTGC GGTGCTGGCT 1140
  GCACACTGTC CCTTCTACAG CTGGAAGAGA GTGTTCCTAA CCCACCCTGC CACCTGCTAC 1200
  AGGACCACCT GCCCCAGGCCC CTGTGACTCG CAGCCCTGCC AGAATGGAGG CACACATGTGTT 1260
  CCAGAAGGAC TGGACCGGCTA CCAGTGCCTC TGCCCGCTGG CCTTTGGAGG GGAGGCTAAC 1320
  TGTGCCCTGA AGCTGAGCCT GGAATGCAGG GTCGGACCTCC TCTTCCTGCT GGACAGCTCT 1380
  GCGGGCACCA CTCTGGACGG CTTCCTGCGG GCCAAAGTCT TCGTGAAGCG GTTTGTGCGG 1440
  GCCGTGCTGA GCGAGGACTC TCGGGCCCGA GTGGGTGTGG CCACATACAG CAGGGAGCTG 1500
  CTGGTGGCGG TGCCTGTGGG GGAGTACCAG GATGTGCCTG ACCTGGTCTG GAGCCTCGAT 1560
  GGCATTCCCT TCCGTGGTGG CCCCACCCTG ACGGGCAGTG CCTTGCGGCA GGCGGCAGAG 1620
  CGTGGCTTCG GGAGCGCCAC CAGGACAGGC CAGGACCGGC CACGTAGAGT GGTTGGTTTTG 1680
  CTCACTGAGT CACACTCCGA GGATGAGGTT GCGGGCCCAG CGCGTCACGC AAGGGCGCGA 1740
  GAGCTGCTCC TGCTGGGTGT AGGCAGTGAG GCCGTGCGGG CAGAGCTGGA GGAGATCACA 1800
  GGCAGCCCAA AGCATGTGAT GGTCTACTCG GATCCTCAGG ATCTGTTCAA CCAAATCCCT 1860
  GAGCTGCAGG GGAAGCTGTG CAGCCGGCAG CGGCCAGGGT GCCGGACACA AGCCCTGGAC 1920
  CTCGTCTTCA TGTTGGACAC CTCTGCCTCA GTAGGGCCCG AGAATTTTGC TCAGATGCAG 1980
  AGCTTTGTGA GAAGCTGTGC CCTCCAGTTT GAGGTGAACC CTGACGTGAC ACAGGTCGGC 2040
  CTGGTGGTGT ATGGCAGCCA GGTGCAGACT GCCTTCGGGC TGGACACCAA ACCCACCCGG 2100
  GCTGCGATGC TGCGGGCCAT TAGCCAGGCC CCCTACCTAG GTGGGGTGGG CTCAGCCGGC 2160
  ACCGCCCTGC TGCACATCTA TGACAAAGTG ATGACCGTCC AGAGGGGTGC CCGGCCTGGT 2220
  GTCCCCAAAG CTGTGGTGGT GCTCACAGGC GGGAGAGGCG CAGAGGATGC AGCCGTTCCT 2280
  GCCCAGAAGC TGAGGAACAA TGGCATCTCT GTCTTGGTCG TGGGCGTGGG GCCTGTCCTA 2340
  AGTGAGGGTC TGCGGAGGCT TGCAGGTCCC CGGGGATTCC TGATCCACGT GGCAGCTTAC 2400
  GCCGACCTGC GGTACCACCA GGACGTGCTC ATTGAGTGGC TGTGTGGAGA AGCCAAGCAG 2460
  CCAGTCAACC TCTGCAAACC CAGCCCGTGC ATGAATGAGG GCAGCTGCGT CCTGCAGAAT 2520
  GGGAGCTACC GCTGCAAGTG TCGGGATGGC TGGGAGGGCC CCCACTGCGA GAACCGTGAG 2580
  TGGAGCTCTT GCTCTGTATG TGTGAGCCAG GGATGGATTC TTGAGACGCC CCTGAGGCAC 2640
  ATGGCTCCCG TGCAGGAGGG CAGCAGCCGT ACCCCTCCCA GCAACTACAG AGAAGGCCTG 2700
  GGCACTGAAA TGGTGCCTAC CTTCTGGAAT GTCGTGTGCC CAGGTCCTTA GAATGTCTGC 2760
  TTCCCGCCGT GGCCAGGACC ACTATTCTCA CTGAGGGAGG AGGATGTCCC AACTGCAGCC 2820
  ATGCTGCTTA GAGACAAGAA AGCAGCTGAT GTCACCCACA AACGATGTTG TTGAAAAGTT 2880
  TTGATGTGTA AGTAAATACC CACTTTCTGT ACCTGCTGTG CCTTGTTGAG GCTATGTCAT 2940
  CTGCCACCTT TCCCTTGAGG ATAAACAAGG GGTCCTGAAG ACTTAAATTT AGCCGGCCTGA 3000
  CGTTCCTTTG CACACAATCA ATGCTCGCCA GAATGTTGTT GACACAGTAA TGCCCAGCAG 3060
  AGGCCTTTAC TAGAGCATCC TTTGGACGGC GAAGGCCACG GCCTTTCAAG ATGGAAAGCA 3120
  GCAGCTTTTC CACTTCCCCA GAGACATTCT GGATGCATTT GCATTGAGTC TGAAAGGGGG 3180
  CTTGAGGGAC GTTTGTGACT TCTTGGCGAC TGCCTTTTGT GTGTGGAAGA GACTTGGAAA 3240
  GGTCTCAGAC TGAATGTGAC CAATTAACCA GCTTGGTTGA TGATGGGGA GGGGCTGAGT 3300
  TGTGCATGGG CCCAGGTCTG GAGGGCCACG TAAAATCGTT CTGAGTCGTG AGCAGTGTCC 3360
  ACCTTGAAGG TCTTC
```

A165 Protein sequence
Gene name: ESTs
Unigene number: Hs.157601
Protein Accession #:   none found
Signal sequence: 1-17
Transmembrane domains: none found
VGW domains: 49-223; 341-518; 529-706
EGF domains: 298-333; 715-748
Cellular Localization: secreted

```
    1         11        21        31        41        51
    |         |         |         |         |         |
MPPFLLLEAV CVFLFSRVPP SLPLQEVBVS KETIGKISAA SKMMWCSAAV DIMFLLDGSN    60
SVGKGSPERS KHFAITVCDG LDISPERVRV GAFQFSSTPH LEFPLDSFST QQEVKARIKR   120
MVFKGGRTET ELALKYLLHR GLPGGRNASV PQILIIVTDG KSQGDVALPS KQLKERGVTV   180
FAVGVRPPRW EELHALASEP RGQBVLLAEQ VEDATNGLFS TLSSSAICSS ATPDCRVEAH   240
PCEHRTLEAV REFAGNAPCW RGSRRTLAVL AAHCPFYSWK RVFLTHPATC YRTTCPGPCD   300
SQPCQNGGTC VPEGLDGYQC LCPLAFGGEA NCALKLSLEC RVDLLFLLDS SAGTTLDGFL   360
RAKVFVKRFV RAVLSEDSRA RVGVATYSRE LLVAVPVGEY QDVPDLVWSL DGIPFRGGPT   420
LTGSALRQAA ERGFGSATRT GQDRPRRVVV LLTESHSEDE VAGPARHARA RELLLLGVGS   480
EAVRAELEEI TGSPKHVMVY SDPQDLFNQI PELQGKLCSR QRPGCRTQAL DLVFMLDTSA   540
SVGPENFAQM QSPVRSCALQ FEVNPDVTQV GLVVYGSQVQ TAFGLDTKPT RAAMLRAISQ   600
APYLGGVGSA GTALLHIYDK VMTVQRGARP GVPKAVVVLT GGRGAEDAAV PAQKLRNNGI   660
SVLVVGVGPV LSEGLRRLAG PRDSLIHVAA YADLRYHQDV LIEWLCGEAK QFVNLCKPSP   720
CMNEGSCVLQ NGSYRCKCRD GWEGPHCENR EWSSCSVCVS QGWILETPLR HMAPVQEGSS   780
RTPPSNYREG LGTEMVPTPW NVCAPGP
```

A166 DNA sequence
Gene name: Homo sapiens beta-1 adrenergic receptor mRNA, 3' UTR
Unigene number: Hs.37744
Probeset Accession #:   AA011176
Nucleic Acid Accession #:   AF272890
Coding Sequence: 87-1520 (underlined sequences correspond to start and stop codons)

```
    1         11        21        31        41        51
    |         |         |         |         |         |
TGCTACCCGC GCCCGGGCTT CTGGGGTGTT CCCCAACCAC GGCCCAGCCC TGCCACACCC    60
CCCGCCCCTG GCCTCCGCAG CTCGGCATGG GCGCGGGGGT GCTCGTCCTG GGCGCCTCCG   120
AGCCCGGTAA CCTGTCGTCG GCCGCCACCGC TCCCCGACGG CGGCGCCACC AGCCCCGAGC   180
TGCTGGTGCC CGCGTCGCCG CCCGCCTCGT TGCTGCCTCC CGCCAGCGAA AGCCCCGAGC   240
CGCTGTCTCA GCAGTGGACA GCGGGCATGG GTCTGCTGAT GGCGCTCATC GTGCTGCTCA   300
TCGTGGCGGG CAATGTGCTG GTGATCGTGG CCATCGCCAA GACGCCGCGG CTGCAGACGC   360
TCACCAACCT CTTCATCATG TCCCTGGCCA GCGCCGACCT GGTCATGGGG CTGCTGGTGG   420
TGCCGGTCGG GGCCACCATC GTGGTGTGGG GCCGCTGGGA GTACGGCTCC TTCTTCTGCG   480
AGCTGTGGAC CTCAGTGGAC GTGCTGTGCG TGACGGCCAG CATCGAGACC CTGTGTGTCA   540
TTGCCCTGGA CCGCTACCTC GCCATCACCT CGCCCTTCCG CTACCAGAGC CTGCTGACGC   600
GCGCGCGGGC GCGGGGCCTC GTGTGCACCG TGTGGGCCAT CTCGGCCCTG GTGTCCTTCC   660
TGCCCATCCT CATGCACTGG TGGCGGGCGG AGAGCGACGA GGCGCGCCGC TGCTACAACG   720
ACCCCAAGTG CTGCGACTTC GTCACCAACC GGGCCTACGC CATCGCCTCG TCGGTAGTCT   780
CCTTCTACGT GCCCCTGTGC ATCATGGCCT TCGTGTACCT GCGGGTGTTC CGCGAGGCCC   840
AGAAGCAGGT GAAGAAGATC GACAGCTGCG AGCGCCGTTT CCTCGGCGGC CCAGCGCGGC   900
CGCCCTCGCC CTCGCCCTCG CCCGTCCCCG CGCCCGGCGG GCCGCCGGA CCCCCGGCCC   960
CCGCCGCCGC CGCCGCCACC GCCCCGGCTG CCAACGGGCG TGCGGGTAAG CGGCGGCCCT  1020
CGGCGCCTGGT GGCCCTACGC GAGCAGAAGG CGCTCAAGAC GCTGGGCATC ATCATGGGCG  1080
TCTTCACGCT CTGCTGGCTG CCCTTCTTCC TGGCCAACGT GGTGAAGGCC TTCCACCGCG  1140
AGCTGGTGCC CGACCGCCTC TTCGTCTTCT TCAACTGGCT GGGCTACGCC AACTCGGCCT  1200
TCAACCCCAT CATCTACTGC CGCAGCCCCG ACTTCCGCAA GGCCTTCCAG GGACTGCTCT  1260
GCTGCGCGCG CAGGGCTGCC CGCCGGCGCC ACGCGACCCA CGGAGACCGG CCGCGGCCT  1320
CGGGCTGTCT GGCCCGGCCC GGACCCCTGC CATCGCCCGG GGCCGCCTCG GACGACGACG  1380
ACGACGATGT CGTCGGGGCC ACGCCGCCCG CGCGCCTGCT GGAGCCCTGG GCCGGCTGCA  1440
ACGGCGGGGAC GGCGGCGGAC AGCGACTCGA GCCTGGACGA GCCGTGCCGC CCCGGCTTCG  1500
CCTCGGAATC CAAGGTGTAG GGCCCGGCGC GGGGCGCGGA CTCCGGGCAC GGCTTCCCAG  1560
GGGAACGAGG AGATCTGTGT TTACTTAAGA CCGATAGCAG GTGAACTCGA AGCCCACAAT  1620
CCTCGTCTGA ATCATCCGAG GCAAAGAGAA AAGCCACGGA CCGTTGCACA AAAAGGAAAG  1680
TTTGGGAAGG GATGGGAGAG TGGCTTGCTG ATGTTCCTTG TTG
```

A167 Protein sequence
Gene name: Homo sapiens beta-1 adrenergic receptor mRNA, 3' UTR
Unigene number: Hs.37744
Protein Accession #:   AA011176
Signal sequence: none found
Transmembrane domains: 62-84, 95-117, 135-157, 177-198, 226-248
Pfam domain: 7tm_1 [75-377]
Cellular Localization: plasma membrane

```
    1         11        21        31        41        51
    |         |         |         |         |         |
MGAGVLVLGA SEPGNLSSAA PLPDGAATAA RLLVPASPPA SLLPPASESP EPLSQQWTAG    60
MGLLMALIVL LIVAGNVLVI VAIAKTPRLQ TLTNLFIMSL ASADLVMGLL VVPPGATIVV   120
```

910

```
WGRWEYGSFF  CELWTSVDVL  CVTASIETLC  VIALDRYLAI  TSPFRYQSLL  TRARARGLVC  180
TVWAISALVS  FLPILMHWWR  AESDEARRCY  NDPKCCDFVT  NRAYAIASSV  VSFYVPLCIM  240
AFVYLRVFRE  AQKQVKKIDS  CERRFLGGPA  RPPSPSPSPV  PAPAPPPGPP  RPAAAAATAP  300
LANGRAGKRR  PSRLVALREQ  KALKTLGIIM  GVFTLCWLPF  FLANVVKAFH  RELVPDRLFV  360
FFNWLGYANS  AFNPIIYCRS  PDFRKAFQGL  LCCARRAARR  RHATHGDRPR  ASGCLARPGP  420
PPSPGAASDD  DDDDVVGATP  PARLLEPWAG  CNGGAAADSD  SSLDEPCRPG  FASESKV
```

<u>A168 DNA sequence</u>
Gene name: CEGP1
Unigene number: Hs.222399
Probeset Accession #:   AA256485
Nucleic Acid Accession #:   AJ400877
Coding sequence: 81-3080 (underlined sequences correspond to start and stop codons)

```
1           11          21          31          41          51
|           |           |           |           |           |
GGCGTCCGCG  CACACCTCCC  CGCGCCGCCG  CCGCCACCGC  CCGCACTCCG  CCGCCTCTGC   60
CCGCAACCGC  TGAGCCATCC  ATGGGGGTCG  CGGGCCGCAA  CCGTCCGGGG  GCGGCCTGGG  120
CGGTGCTGCT  GCTGCTGCTG  CTGCTGCCGC  CACTGCTGCT  GCTGGCGGGG  GCCGTCCCGC  180
CGGGTCGGGG  CCGTGCCGCG  GGGCCGCAGG  AGGATGTAGA  TGAGTGTGCC  CAAGGGCTAG  240
ATGACTGCCA  TGCCGACGCC  CTGTGTCAGA  ACACACCCAC  CTCCTACAAG  TGCTCCTGCA  300
AGCCTGGCTA  CCAAGGGGAA  GGCAGGCAGT  GTGAGGACAT  CGATGAATGT  GGAAATGAGC  360
TCAATGGAGG  CTGTGTCCAT  GACTGTTTGA  ATATTCCAGG  CAATTATCGT  TGCACTTGTT  420
TTGATGGCTT  CATGTTGGCT  CATGACGGTC  ATAATTGTCT  TGATGTGGAC  GAGTGCCTGG  480
AGAACAATGG  CGGCTGCCAG  CATACCTGTG  TCAACGTCAT  GGGGAGCTAT  GAGTGCTGCT  540
GCAAGGAGGG  GTTTTTCCTG  AGTGACAATC  AGCACACCTG  CATTCACCGC  TCGGAAGAGG  600
GCCTGAGCTG  CATGAATAAG  GATCACGGCT  GTAGTCACAT  CTGCAGGGAG  GCCCCAAGGG  660
GCAGCGTCGC  CTGTGAGTGC  AGGCCTGGTT  TTGAGCTGGC  CAAGAACCAG  AGAGACTGCA  720
TCTTGACCTG  TAACCATGGG  AACGGTGGGT  GCCAGCACTC  CTGTGACGAT  ACAGCCGATG  780
GCCCAGAGTG  CAGCTGCCAT  CCACAGTACA  AGATGCACAC  AGATGGGAGG  AGCTGCCTTG  840
AGCGAGAGGA  CACTGTCCTG  GAGGTGACAG  AGAGCAACAC  CACATCAGTG  GTGGATGGGG  900
ATAAACGGGT  GAAACGGCGG  CTGCTCATGG  AAACGTGTGC  TGTCAACAAT  GGAGGCTGTG  960
ACCGACCTG  TAAGGATACT  TCGACAGGTG  TCCACTGCAG  TTGTCCTGTT  GGATTCACTC  1020
TCCAGTTGGA  TGGGAAGACA  TGTAAAGATA  TTGATGAGTG  CCAGACCCGC  AATGGAGGTT  1080
GTGATCATTT  CTGCAAAAAC  ATCGTGGGCA  GTTTTGACTG  CGGCTGCAAG  AAAGGATTTA  1140
AATTATTAAC  AGATGAGAAG  TCTTGCCAAG  ATGTGGATGA  GTGCTCTTTG  GATAGGACCT  1200
GTGACCACAG  CTGCATCAAC  CACCCTGGCA  CATTTGCTTG  TGCTTGCAAC  CGAGGGTACA  1260
CCCTGTATGG  CTTCACCCAC  TGTGGAGACA  CCAATGAGTG  CAGCATCAAC  AACGGAGGCT  1320
GTCAGCAGGT  CTGTGTGAAC  ACAGTGGGCA  GCTATGAATG  CCAGTGCCAC  CCTGGGTACA  1380
AGCTCCACTG  GAATAAAAAA  GACTGTGTGG  AAGTGAAGGG  GCTCCTGCCC  ACAAGTGTGT  1440
CACCCCGTGT  GTCCCTGCAC  TGCGGTAAGA  GTGGTGGAGG  AGACGGGTGC  TTCCTCAGAT  1500
GTCACTCTGG  CATTCACCTC  TCTTCAGATG  TCACCACCAT  CAGGACAAGT  GTAACCTTTA  1560
AGCTAAATGA  AGGCAAGTGT  AGTTTGAAAA  ATGCTGAGCT  GTTTCCCGAG  GGTCTGCGAC  1620
CAGCACTACC  AGAGAAGCAC  AGCTCAGTAA  AAGAGAGCTT  CCGCTACGTA  AACCTTACAT  1680
GCAGCTCTGG  CAAGCAAGTC  CCAGGAGCCC  CTGGCCGACC  AAGCCACCCCT  AAGGAAATGT  1740
TTATCACTGT  TGAGTTTGAG  CTTGAAACTA  ACCAAAAGGA  GGTGACAGCT  TCTTGTGACC  1800
TGAGCTGCAT  CGTAAAGCGA  ACCGAGAAGC  GGCTCCGTAA  AGCCATCCGC  ACGCTCAGAA  1860
AGGCCGTCCA  CAGGGAGCAG  TTTCACCTCC  AGCTCTCAGG  CATGAACCTC  GACGTGGCTA  1920
AAAAGCCTCC  CAGAACATCT  GAACGCCAGG  CAGAGTCCTG  TGGAGTGGGC  CAGGGTCATG  1980
CAGAAAACCA  ATGTGTCAGT  TGCAGGGCTG  GGACCTATTA  TGATGGAGCA  CGAGAACGCT  2040
GCATTTTATG  TCCAAATGGA  ACCTTCCAAA  ATGAGGAAGG  ACAAATGACT  TGTGAACCAT  2100
GCCCAAGACC  AGGAAATTCT  GGGGCCCTGA  AGACCCCAGA  AGCTTGGAAT  ATGTCTGAAT  2160
GTGGAGGTCT  GTGTCAACCT  GGTGAATATT  CTGCAGATGG  CTTTGCACCT  TGCCAGCTCT  2220
GTGCCCTGGG  CACGTTCCAG  CCTGAAGCTG  GTCGAACTTC  CTGCTTCCCC  TGTGGAGGAG  2280
GCCTTGCCAC  CAAACATCAG  GGAGCTACTT  CCTTCAGGA  CTGTGAAACC  AGAGTTCAAT  2340
GTTCACCTGG  ACATTTCTAC  AACACCACCA  CTCACCGATG  TATTCGTTGC  CCAGTGGGAA  2400
CATACCAGCC  TGAATTTGGA  AAAAATAATT  GTGTTTCTTG  CCCAGGAAAT  ACTACGACTG  2460
ACTTTGATGG  CTCCACAAAC  ATAACCCAGT  GTAAAAACAG  AAGGATGTGGA  GGGGAGCTGG  2520
GAGATTTCAC  TGGGTACATT  GAATCCCCAA  ACTACCCAGG  CAATTACCCA  GCCAACACCG  2580
AGTGTACGTG  GACCATCAAC  CCACCCCCCA  AGCGCGCGAT  CCTGATCGTG  GTCCCTGAGA  2640
TCTTCCTGCC  CATAGAGGAC  GACTGTGGGG  ACTATCTGGT  GATGCGGAAA  ACCTCTTCAT  2700
CCAATTCTGT  GACAACATAT  GAAACCTGCC  AGACCTACGA  ACGCCCCATC  GCCTTCACCT  2760
CCAGGTCAAA  GAAGCTGTGG  ATTCAGTTCA  AGTCCAATGA  AGGGAACCAGC  GCTAGAGGGT  2820
TCCAGGTCCC  ATACGTGACA  TATGATGAGG  ACTACCAGGA  ACTCATTGAA  GACATAGTTC  2880
GAGATGGCAG  GCTCTATGCA  TCTGAGAACC  ATCAGGAAAT  ACTTAAGGAT  AAGAAACTTA  2940
TCAAGGCTGT  GTTTGATGTC  CTGGCCCATC  CCCAGAACTA  TTTCAAGTAC  ACAGCCCAGG  3000
AGTCCCGAGA  GATGTTTCCA  AGATCGTTCA  TCCGATTGCT  ACGTTCCAAA  GTGTCCAGGT  3060
TTTTGAGACC  TTACAAATGA  CTCAGCCAC  GTGCCACTCA  ATACAAATGT  TCTGCTATAG  3120
GGTTGGTGGG  ACAGAGCTGT  CTTCCTTCTG  CATGTCAGCA  CAGTCGGGTA  TTGCTGCCTC  3180
CCGTATCAGT  GACTCATTAG  AGTTCAATTT  TTATAGATAA  TACAGATATT  TTGGTAAATT  3240
GAACTTGGTT  TTTCTTTCCC  AGCATCGTGG  ATGTAGACTG  AGAATGGCTT  TGAGTGCAT  3300
CAGCTTCTCA  CTGCTGTGGG  CGGATGTCTT  GGATAGATCA  CGGGCTGGCT  GAGCTGGACT  3360
TTGGTCAGCC  TAGGTGAGAC  TCACCTGTCC  TTCTGGGGTC  TTACTCCTCC  TCAAGGAGTC  3420
TGTAGTGGAA  AGGAGGCCAC  AGAATAAGCT  GCTTATTCTG  AAACTTCAGC  TTCCTCTAGC  3480
CCGGCCCTCT  CTAAGGGAGC  CCTCTGCACT  CGTGTGCAGG  CTCTGACCAG  GCAGAACAGG  3540
CAAGAGGGGA  GGGAAGGAGA  CCCCTGCAGG  CTCCCTCCAC  CCACCTTGAG  ACCTGGGAGG  3600
ACTCAGTTTC  TCCACAGCCT  TCTCCAGCCT  GTGTGATACA  AGTTTGATCC  CAGGAACTTG  3660
AGTTCTAAGC  AGTGCTCGTG  AAAAAAAAAA  GCAGAAAGAA  TTAGAAATAA  ATAAAAACTA  3720
AGCACTTCTG  GAGACAT
```

<u>A169 Protein sequence</u>
Gene name: CEGP1
Unigene number: Hs.222399

```
Probeset Accession #:              AA256485
Protein Accession #:               CAB92285
Signal sequence:                   1-31
Transmembrane domains:             none
PPAM domains:      EGF-like_domains [49-84,132-167,177-213,286-321,407-442]  CUB_domain [809-918]
Cellular Localization:                        may be secreted
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MGVAGRNRPG AAWAVLLLLL LLPPLLLLAG AVPPGRGRAA GPQEDVDECA QGLDDCHADA    60
LCQNTPTSYK CSCKPGYQGE GRQCEDIDEC GNELNGGCVH DCLNIPGNYR CTCFDGPMLA   120
HDGHNCLDVD ECLENNGGCQ HTCVNVMGSY ECCCKEGFFL SDNQHTCIHR SEEGLSCMNK   180
DHGCSHICKE APRGSVACEC RPGFELAKNQ RDCILTCNHG NGGCQHSCDD TADGPECSCH   240
PQYKMETDGR SCLEREDTVL EVTESNTTSV VDGDKRVKRR LLMETCAVNN GGCDRTCKDT   300
STGVHCSCPV GFTLQLDGKT CKDIDECQTR NGGCDHFCKN IVGSFDCGCK KGFKLLTDEK   360
SCQDVDECSL DRTCDHSCIN HPGTFACACN RGYTLYGFTH CGDTNECSIN NGGCQQVCVN   420
TVGSYECQCH PGYKLHWNKK DCVEVKGLLP TSVSPRVSLH CGKSGGGDGC FLRCHSGIHL   480
SSDVTTIRTS VTFKLNEGKC SLKNAELFPE GLRPALPEKH SSVKESFRYV NLTCSSGKQV   540
PGAPGRPSTP KEMFITVEFE LETNQKEVTA SCDLSCIVKR TEKRLRKAIR TLRKAVHREQ   600
PHLQLSGMNL DVAKKPPRTS ERQAESCGVG QGHAENQCVS CRAGTYYDGA RERCILCPNG   660
TFQNEEGQMT CEPCPRPGNS GALKTPEAWN MSECGGLCQP GEYSADGFAP CQLCALGTFQ   720
FEAGRTSCFP CGGGLATKHQ GATSFQDCET RVQCSPGHFY NTTTHRCIRC PVGTYQPEFG   780
KNNCVSCPGN TTTDFDGSTN ITQCKNRRCG GELGDFTGYI ESPNYPGNYP ANTECTWTIN   840
PPPKRRILIV VPEIFLPIED DCGDYLVMRK TSSSNSVTTY ETCQTYERPI AFTSRSKKLW   900
IQFKSNEGNS ARGFQVPYVT YDEDYQELIE DIVRDGRLYA SENRQEILKD KKLIKALPDV   960
LAHPQNYFKY TAQESREMFP RSFIRLLRSK VSRFLRPYK
```

A170 DNA sequence

```
Gene name:                    DEME-6 protein (KIAA0452)
Unigene number:               Hs.125783
Probeset Accession #:         AL039402
Nucleic Acid Accession #:     AF007170
Coding sequence:              1-1725 (underlined sequences correspond to stop codon)
```

```
1          11         21         31         41         51
|          |          |          |          |          |
AAGGAGGCGG CCTCCGGGAA AAGCGACCGC AGGACTCCTG AGAGCAGCCT CCATGAGGCC    60
CTGGACCAGT GCATGACCGC CCTGGACCTC TTCCTCACCA ACCAGTTCTC AGAAGCACTC   120
AGCTACCTCA AGCCCAGAAC CAAGGAAAGC ATGTACCACT CACTGACATA TGCCACCATC   180
CTGGAGATGC AGGCCATGAT GACCTTTGAC CCTCAGGACA TCCTGCTTGC CGGCAACATG   240
ATGAAGGAGG CACAGATGCT GTGTCAGAGG CACCGGAGGA AGTCTTCTGT AACAGATTCC   300
TTCAGCAGCC TGGTGAACCG CCCCCACGCTG GGCCAATTCA CTGAAGAAGA AATCCACGCT   360
GAGGTCTGCT ATGCAGAGTG CCTGCTGCAG CGAGCAGCCC TGACCTTCCT GCAGGACGAG   420
AACATGGTGA GCTTCATCAA AGGCGGCATC AAAGTTCGAA ACAGCTACCA GACCTACAAG   480
GAGCTGGACA GCCTTGTTCA GTCCTCACAA TACTGCAAGG GTGAGAACCA CCCGCACTTT   540
GAAGGAGGAG TGAAGCTTGG TGTAGGGGCC TTCAACCTGA CACTGTCCAT GCTTCCTACT   600
AGGATCCTGA GGCTGTTGGA GTTTGTGGGG TTTTGCAGGA ACAAGGACTA TGGGCTGCTG   660
CAGCTGGAGG AGGGAGCGTC AGGGCACAGC TTCCGCTCTG TGCTCTGTGT CATGCTCCTG   720
CTGTGCTACC ACACCTTCCT CACCTTCGTG CTCGGTACTG GGAACGTCAA CATCGAGGAG   780
GCCGAGAAGC TCTTGAAGCC CTACCTGAAC CGGTACCCTA AGGGTGCCAT CTTCCTGTTC   840
TTTGCAGGGA GGATTGAAGT CATTAAAGGC AACATTGATG CAGCCATCCG GCGTTTCTAG   900
GAGTGCTGTG AGGCCCAGCA GCACTGGAAG CAGTTTCACC ACATGTGCTA CTGGGAGCTG   960
ATGTGGTGCT TCACCTACAA GGGCCAGTGG AAGATGTCCT ACTTCTACGC CGACCTGCTC   1020
AGCAAGGAGA ACTGCTGGTC CAAGGCCACC TACATTTACA TGAAGGCCGC CTACCTCGAA   1080
ATGTTTGGGA AGGAGGACCA CAAGCCGTTC GGGGACGACG AAGTGGAATT ATTTCGAGCT   1140
GTGCCAGGCC TGAAGCTCAA GATTGCTGGG AAATCTCTAC CCACAGAGAA GTTTGCCATC   1200
CGGAAGTCCC GGCGCTACTT CTCCTCCAAC CCTATCTCGC TGCCAGTGCC TGCTCTGGAA   1260
ATGATGTACA TCTGGAACGG CTACGCCGTG ATTGGGAAGC AGCCGAAACT CACGGATGGG   1320
ATACTTGAGA TTATCACTAA GGCTGAAGAG ATGCTGGAGA AAGGCCCAGA GAACGAGTAC   1380
TCAGTGGATG ACGAGTGCTT GGTGAAATTG TTGAAAGGCC TGTGTCTGAA ATACCTGGGC   1440
CGTGTCCAGG AGGCCGAGGA GAATTTTAGG AGCATCTCTG CCAATGAAAA GAAGATTAAA   1500
TATGACCACT ACTTGATCCC AAACGCCCTG CTGGAGCTGG CCCTGCTGCT TATGGAGCAA   1560
GACAGAAACG AAGAGGCCAT CAAACTTTTG GAATCTGCCA AGCAAAACTA CAAGAATTAC   1620
TCCATGGAGT CAAGGACACA CTTTCGAATC CAGGCAGCCA CACTCCAAGC CAAGTCTTCC   1680
CTAGAGAACA GCAGCAGATC CATGGTCTCA TCAGTGTCCT TGTAGCTTTG TGCAGCAGTT   1740
CCGGGCTGGA AGACAGAGAC AGCTGGACAG AGCTCCTGAA AACATTTCAA AATACCCCCT   1800
CCCCCTGCCC TGCCCTGCCT TTGGGGTCCA CCGGCACTCC AGTTGGATGG CACAACATAG   1860
TGTATCCGTG CAGAAGCCGA GCTGGCATTT TCACCAGTGT AGCCAAGGGC CTTGCCAAG    1920
GGCAGAGCAG GTGGAGCCCT CTGCCTGCCC TATCACACAT ACGGGTACTT GCTTTTCACT   1980
GTGATGTTTA AGAGAATGTA TGAACAGTTT ACATTTTCCT TAGAAATACA TTGATGGGAT   2040
CACAGTTGGC TTTAAAAACC AACAACAATC AACCACCTGT AAGTCTTTGT CTTCACCTAT   2100
TATCATCTGG AGGTAAATCT CTTTATATGA TGATGCCAAA GGGCAAATTG CTTTTCAAAT   2160
TCAGCAAGTT CTCAGCTTGT GTGACGGAAG GTCCTTCAGA GGACCTGAGG AATGCCTGGG   2220
AGAGGCTAAG CCTCAGGCTT CAATGCTTCT GGGGTTGGGC ATGAGGATGT ACACAGACAC   2280
CCACTACCTT ACTACTCACA CTTCATTTCA CTCCTTTTGT AAATTTCCAA TTTAAAAATC   2340
AAGCACGTCT TTTTAGTGAG ATAAAATCTG AGCTCTTCTG TAGAAAAATC AATCTCTACC   2400
AGTAGAAAAT GCCAGGGCTT GATGGAAGAG CTGTGTAGCC CTTTCTATGC CAAAGCCAGG   2460
AAATTTGGGG GGCAGGAGGA GGTTCTCAGA ATCCAGTCTG TATCTTTGCT GTATGCCAAA   2520
CTGAAACCAC TGGGAATAAT TTATGAAACA TAAAAATCTT CTGTACTTCA CTCCAAGGTA   2580
CATTTGCTTA CTGACAGCAT TTTTGTTAAA ACTGTTATTC TTGAAAAAAA AAAAAAAAAA   2640
AA
```

A171 Protein sequence

Gene name: DBME-6 protein (KIAA0452)
Unigene number: Hs.125783
Probeset Accession #:   AL039402
Protein Accession #:   AAC39582
Signal sequence: none
Transmembrane domains: 210-226
Cellular Localization: plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
MTALDLPLTN QFSEALSYLK PRTKESMYHS LTYATILEMQ AMMTFDPQDI LLAGNMMKEA   60
QMLCQRERRK SSVTDSFSSL VNRPTLGQFT EEEIHAEVCY AECLLQRAAL TFLQDENMVS  120
FIKGGIKVRN SYQTYKELDS LVQSSQYCKG ENEPHFEGGV KLGVGAFNLT LSMLPTHILE  180
LLEFVGFSGN KDYGLLQLEE GASGHSFRSV LCVMLLLCYH TFLTFVLGTG NVNIEEAEKL  240
LKPYLNRYPK GAIFLFFAGR IEVIKGNIDA AIRRFEECCE AQQHWKQFHH MCYWELMWCF  300
TYKGQWKMSY FYADLLSKEN CWSKATYIYM KAAYLSMFGK EDHKPFGDDE VELFRAVPGL  360
KLKIAGKSLP TEKFAIRKSR RYFSSNPISL PVPALEMMYI WNGYAVIGKQ PKLTDGILEI  420
ITKAEEMLEK GPENEYSVDD ECLVKLLKGL CLKYLGRVQE AEENFRSISA NEKKIKYDHY  480
LIPNALLELA LLLMEQDRNE EAIKLLESAK QNYKNYSMES RTHFRIQAAT LQAKSSLENS  540
SRSMVSSVSL
```

<u>A172 DNA sequence</u>
Gene name: EST
Unigene number: Hs.200102
Probeset Accession #:   AL117406
Nucleic Acid Accession #:   none found
Coding sequence: 1-4044 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGACTAGGA AGAGGACATA CTGGGTGCCC AACTCTTCTG GTGGCCTCGT GAATCGTGGC   60
ATCGACATAG GCGATGACAT GGTTTCAGGA CTTATTTATA AAACCTATAC TCTCCAAGAT  120
GGCCCCTGGA GTCAGCAAGA GAGAAATCCT GAGGCTCCAG GGAGGGCAGC TGTCCCACCG  180
TGGGGGAAGT ATGATGCTGC CTTGAGAACC ATGATTCCCT TCCGTCCCAA GCCGAGGTTT  240
CCTGCCCCCC AGCCCCTGGA CAATGCTGGC CTGTTCTCCT ACCTCACCGT GTCATGGCTC  300
ACCCCGCTCA TGATCCAAAG CTTACGGAGT CGCTTAGATG AGAACACCAT CCCTCCACTG  360
TCAGTCCATG ATGCCTCAGA CAAAAATGTC CAAAGGCTTG CCGGCCTTTG GGAGAAGAA   420
GTCTCAAGGC GAGGGGATTGA AAAAGCTTCA GTGCTTCTGG TGATGCTGAG GTTCCAGAGA  480
ACAAGGTTGA TTTTCGATGC ACTTCTGGGG ATCTGCTTCT GCATTGCCAG TGTACTCGGG  540
CCAATATTGA TTATACCAAA GATCCTGGAA TATTCAGAAG AGCAGTTGGG GAATGTTGTC  600
CATGGAGTGG GACTCTGCTT TGCCCTTTTT CTCTCCGAAT GTGTGAAGTC TCTGAGTTTC  660
TCCTCCAGTT GGATCATCAA CCAACGCACA GCCATCAGGT TCCGAGCAGC TGTTTCCTCC  720
TTTGCCTTTG AGAAGCTCAT CCAAATTAAG TCTGTAACTAC ACATCACCTC AGGAGAGGCC  780
ATCAGCTTCT TCACCGGTGA TGTAAACTAC CTGTTTGAAG GGGTGTGCTA TGGACCCCTA  840
GTACTGATCA CCTGCGCATC GCTGGTCATC TGCAGCATTT CTTCCTACTT CATTATTGGA  900
TACACTGCAT TTATTGCCAT CTTATGCTAT CTCCCTGGTTT TCCCACTGGC GGTATTCATG  960
ACAAGAATGG CTGTGAAGGC TCAGCATCAC ACATCTGAGG TCAGCGACCA GCGCATCCGT 1020
GTGACCAGTG AAGTTCTCAC TTGCATTAAG CTGATTAAAA TGTACACATG GGAGAAACCA 1080
TTTGCAAAAA TCATTGAAGG TATGGAAAGT CTGACTTTCT GCTCCAAACC TGGTGATGGC 1140
ATGGCCTTCA GCATGCTGGC CTCCTTGAAT CTCCTTCGGC TGTCAGTGTT CTTTGTGCCT 1200
ATTCCAGTCA AAGGTCTCAC GAATTCCAAG TCTGCAGTGA TGAGGTTCAA GAAGTTTTTC 1260
CTCCAGGAGA GCCCTGTTTT CTATGTCCAG ACATTACAAG ACCCCAGCAA AGCTCTGGTC 1320
TTTGAGGAGG CCACCTTGTC ATGGCAACAG ACCTGTCCCG GGAGTGTCAA TGGGGCACTG 1380
GAGCTGGAGA GGAACGGGCA TGCTTCTGAG GGGATGACCA GGCCTAGAGA TGCCCTCGGG 1440
CCAGAGGAAG AAGGGAACAG CCTGGGCCCA GAGTTGCACA AGATCAACCT GGTGGTGTCC 1500
AAGGGGATGA TGTTAGGGGT CTGCGGCAAC ACGGGGAGTG GTAAGAGCAG CCTGTTGTCA 1560
GCCATCCTGG AGGAGATGCA CTTGCTCGAG GGCTCGGTGG GGGTGCAGGG AAGCCTGGCC 1620
TATGTCCCCC AGCAGGCCTG GATCGTCAGC GGGAACATCA GGGAGAACAT CCTCATGGGA 1680
GGCGCATATG ACAAGGCCCG ATACCTCCAG GTGCTCCACT GCTGCTCCCT GAATGCGGAC 1740
CTGGAACTTC TGCCCGTTGG AGACATGACA GAGATTGGAG AGCGGGGGCCT CAACCTCTCT 1800
GGGGGGCAGA AACAGAGGAT CAGCCTGGCC CGCGCCGTCT ATTCCGACCG TCAGATCTAC 1860
CTGCTGGACG ACCCCCTGTC TGCTGTGGAC GCCCACGTGG GGAACGCACAT TTTTGAGGAG 1920
TGCATTAAGA AGACACTCAG GGGGAAGACG GTCGTCCTGG TGACCCACCA GCTGCAGTAC 1980
TTAGAATTTT GTGGCCAGAT CATTTTGTTG GAAAATGGGA AAATCTGTGA AAATGGAACT 2040
CACAGTGAGT TAATGCAGAA AAAGGGGAAA TATGCCCAAC TTATCCAGAA GATGCACAAG 2100
GAAGCCACTT CGGACATGTT GCAGGACACA GCAAAGATAG CAGACAAGCC AAAGGTAGAA 2160
AGTCAGGCTC TGGCCACCTC CCTGGAAGAG TCTCTCAACG GAAATGCTGT GCCGGAGCAT 2220
CAGCTCACAC AGGAGGAGGA GATGGAAGAA GGCTCCTTGA GTTGGAGGGT CTACCACCAC 2280
TACATCCAGG CAGCTGGAGG TTACATGGTC TCTTGCATAA TTTTCTTCTT CGTGGTGCTG 2340
ATCGTCTTCT TAACGATCTT CAGCTTCTGG TGGCTGAGCT ACTGGTTGGA GCAGGGCTCG 2400
GGGACCAATA GCAGCCGAGA GAGCAATGGA ACCATGGCAG ACCTGGGCAA CATTGCAGAC 2460
AATCCTCAAC TGTCCTTCTA CCAGCTGGTG TACGGGCTCA ACGCCCTGCT CCTCATCTGT 2520
GTGGGGGTCT GCTCCTCAGG GATTTTCACC AAAGTCACGA GGAAGGCATC CACGGCCCTG 2580
CACAACAAGC TCTTCAACAA GGTTTTCCGC TGCCCCATGA GTTTCTTTGA CACCATCCCA 2640
ATAGGCCGGC TTTTGAACTG CTTCGCAGGG GACTTGGAAC AGCTGGACCA GCTCTTGCCC 2700
ATCTTTTCAG AGCAGTTCCT GGTCCTGTCC TTAATGGTGA TGCCGTCCT GTTGATTGTC 2760
AGTGTGCTGT CTCCATATAT CCTGTTAATG GGAGCCATAA TCATGGTTAT TTGCTTCATT 2820
TATTATATGA TGTTCAAGAA GGCCATCGGT GTGTTCAAGA CTATAGCCGG 2880
TCTCCTTTAT TCTCCCACAT CCTCAATTCT CTGCAAGGCC TGAGCTCCAT CCATGTCTAT 2940
GGAAAAACTG AAGACTTCAT CAGCCAGTTT AAGAGGCTGA CTGATGCGCA GAATAACTAC 3000
CTGCTGTTGT TTCTATCTTC CACACGATGG ATGGCATTGA GGCTGAGAGT CATGACCAAC 3060
CTTGTGACCT TGGCTGTTGC CCTGTTCGTG GCTTTTGGCA TTTCCTCCAC CCCCTACTCC 3120
```

```
TTTAAAGTCA  TGGCTGTCAA  CATCGTGCTG  CAGCTGGCGT  CCAGCTTCCA  GGCCACTGCC  3180
CGGATTGGCT  TGGAGACAGA  GGCACAGTTC  ACGGCTGTAG  AGAGGATACT  GCAGTACATG  3240
AAGATGTGTG  TCTCGGAAGC  TCCTTTACAC  ATGGAAGGCA  CAAGTTGTCC  CCAGGGGTGG  3300
CCACAGCATG  GGGAAATCAT  ATTTCAGGAT  TATCACATGA  AATACAGAGA  CAACACACCC  3360
ACCGTGCTTC  ACGGCATCAA  CCTGACCATC  CGCGGCCACG  AAGTGGTGGG  CATCGTGGGA  3420
AGGACGGGCT  CTGGGAAGTC  CTCCTTGGGC  ATGGCTCTCT  TCCGCCTGGT  GGAGCCCATG  3480
GCAGGCCGGA  TTCTCATTGA  CGGCGTGGAC  ATTTGCAGCA  TCGGCCTGGA  GGACTTGCGG  3540
TCCAAGCTCT  CAGTGATCCC  TCAAGATCCA  GTGCTGCTCT  CAGGAACCAT  CAGATTCAAC  3600
CTAGATCCCT  TTGACCGTCA  CACTGACCAG  CAGATCTGGG  ATGCCTTGGA  GAGGACATTC  3660
CTGACCAAGG  CCATCTCAAA  GTTCCCCAAA  AAGCTGCATA  CAGATGTGGT  GGAAAACGGT  3720
GGAAACTTCT  CTGTGGGGGA  GAGGCAGCTG  CTCTGCATTG  CCAGGGCTGT  GCTTCGCAAC  3780
TCCAAGATCA  TCCTTATCGA  TGAAGCCACA  GCCTCCATTG  ACATGGAGAC  AGACACCCTG  3840
ATCCAGCGCA  CAATCCGTGA  AGCCTTCCAG  GGCTGCACCG  TGCTCGTCAT  TGCCCACCGT  3900
GTCACCACTG  TGCTGAACTG  TGACCACATC  CTGGTTATGG  GCAATGGGAA  GGTGGTAGAA  3960
TTTGATCGGC  CGGAGGTACT  GCGGAAGAAG  CCTGGGTCAT  TGTTCGCAGC  CCTCATGGCC  4020
ACAGCCACTT  CTTCACTGAG  ATAAGGAGAT  GTGGAGACTT  CATGGAGGCT  GGCAGCTGAG  4080
CTCAGAGGTT  CACACAGGTG  CAGCTTCGAG  GCCCACAGTC  TGCGACCTTC  TTGTTTGGAG  4140
ATGAGAACTT  CTCCTGGAAG  CAGGGGTAAA  TGTAGGGGGG  GTGGGGATTG  CTGGATGGAA  4200
ACCCTGGAAT  AGGCTACTTG  ATGGCTCTCA  AGACCTTAGA  ACCCCAGAAC  CATCTAAGAC  4260
ATGGGATTCA  GTGATCATGT  GGTTCTCCTT  TTAACTTACA  TGCTGAATAA  TTTTATAATA  4320
AGGTAAAAGC  TTATAGTTTT  CTGATCTGTG  TTAGAAGTGT  TGCAAATGCT  GTACTGACTT  4380
TGTAAAATAT  AAAACTAAGG  AAAACTCAAA  AAAAAAAAAA  AAAAAAA
```

A173 Protein sequence

Gene name:	EST
Unigene number:	Hs.200102
Probeset Accession #:	AL117406
Protein Accession #:	none found
Signal sequence:	none found
Transmembrane domains:	169-185, 199-215, 275-291, 304-320, 387-403, 770-786, 829-845, 907-923, 927-943,1018-1034
PFAM domains:	ABC_transporter [502-673], ABC_membrane_region [163-432, 771-1060] ATP-binding_domains [508-516, 1139-1147]
Cellular Localization:	plasma membrane

```
1           11          21          31          41          51
|           |           |           |           |           |
MTRKRTYWVP  NSSGGLVNRG  IDIGDDMVSG  LIYKTYTLQD  GPWSQQERNP  EAPGRAAVPP  60
WGKYDAALRT  MIPFRPKPRF  PAPQPLDNAG  LPSYLTVSWL  TPLMIQSLRS  RLDENTIPPL  120
SVHDASDKNV  QRLHRLWEBE  VSRRGIEKAS  VLLVMLRPQR  TRLIFDALLG  ICFCIASVLG  180
PILIIPKILE  YSEEQLGNVV  HGVGLCFALF  LSECVKSLSF  SSSWIINQRT  AIRFRAAVSS  240
FAFEKLIQFK  SVIHITSGEA  ISFFTGDVNY  LFSGVCYGPL  VLITCASLVI  CSISSYFIIG  300
YTAFIAILCY  LLVFPLAVFM  TRMAVKAQHH  TSSVSDQRIR  VTSSVLTCIK  LIKMYTWEKP  360
FAKIIEGMES  LTFCSKPGDG  MAFSMLASLN  LLRLSVFFVP  IAVKGLTNSK  SAVMRFKKFF  420
LQESPVPYVQ  TLQDPSKALV  FEEATLSWQQ  TCPGIVNGAL  ELERNGHASE  GMTRPRDALG  480
PEEEGNSLGP  ELHKINLVVS  KGMMLGVCGN  TGSGKSSLLS  AILEEMHLLE  GSVGVQGSLA  540
YVPQQANIVS  GNIRENILMG  GAYDKARYLQ  VLECCSLNRD  LELLPPGDMT  EIGERGLNLS  600
GGQKQRISLA  RAVYSDRQIY  LLDDPLSAVD  AHVGKHIFEE  CIKKTLRGKT  VVLVTHQLQY  660
LSFCGQIILL  ENGKICENGT  HSELMQKRGK  YAQLIQKMGK  EATSDMLQDT  AKIAEKPKVE  720
SQALATSLEE  SLNGNAVPEH  QLTQEESMEE  GSLSWRVYHH  YIQAAGGYMV  SCIIFPFVVL  780
IVFLTIFSFW  WLSYWLEQGS  GTNSSRESNG  TMADLGNIAD  NPQLSFYQLV  YGLNALLLIC  840
VGVCSSGIFT  KVTRKASTAL  HNKLFNKVFR  CPMSFFDTIP  IGRLLNCFAG  DLEQLDQLLP  900
IFSEQFLVLS  LMVIAVLLIV  SVLSPYIILM  GAIIMVICFI  YYMMFKKAIG  VFKRLENYSR  960
SPLPSHILMS  LQGLSSIHVY  GKTEDFISQF  KRLTDAQWNY  LLLFLSSTRW  MALRLEIMTN  1020
LVTLAVALFV  AFGISSTFYS  FKVMAVNIVL  QLASSFQATA  RIGLETEAQF  TAVERILQYM  1080
KMCVSEAPLH  MEGTSCPQGW  PQHGEIIFQD  YHMKYRDNTP  TVLHGIMLTI  RGHEVVGIVG  1140
RTGSGKSSLG  MALFRLVEPM  AGRILIDGVD  ICSIGLEDLR  SKLSVIPQDP  VLLSGTIRFN  1200
LDPFDRHTDQ  QIWDALERTF  LTKAISKFPK  KLHTDVVENG  GNFSVGERQL  LCIARAVLRN  1260
SKIILIDEAT  ASIDMETDTL  IQRTIREAFQ  GCTVLVIAHR  VTTVLNCDHI  LVMKNGKVVE  1320
FDRPEVLRKK  PGSLFAALMA  TATSSLR
```

A174 DNA sequence

Gene name:	ESTs
Unigene number:	Hs.128899
Probeset Accession #:	AA983251
Nucleic Acid Accession #:	AA983251
Coding sequence:	1-1749 (underlined sequences correspond to start and stop codons)

```
1           11          21          31          41          51
|           |           |           |           |           |
ATGCTGTCTG  GCTTCTTGAT  GAGTCCCAGT  ACCCAGCACA  GAGCACAGTA  CACTCCCGGA  60
GGAAAGAAAC  TTCCGTGGGA  GGCTTCCATC  GGTGCGCACA  CCTCCCGAGG  GCGAGGCAGC  120
GACCGGGAGA  GGGAGAGCCG  GCCGGAGGCT  GCCGGGCTCC  TGTGGGACGG  CGCTGCAGCC  180
GGGGAGGCGG  AGAAGGGGAA  CCGGGGCGAG  CCGCCGCCT  GGATCCGCGC  CCAGCAGCAG  240
CCGCGGCCGC  CGCCAGCTGG  GCAGGCTCCC  GGGACTGCGG  CTGGGGGCGC  GCAGGACCCT  300
CGCCTGCGTC  CTGGACGTTC  CCGGGGGAGG  GTCCGGTTGC  CAGTGAAACC  TCCAGAGGCT  360
TCCGGACGAC  AGCCCCGGGG  GCCTTCTGAC  TGCATCCCGA  GATTTCCATC  AGCGAGTGCA  420
ACTCATAAGG  CAGTCCCTAA  GGGGACCGGG  CCACCGGCTG  AGGACGGGGA  TGGCTTAGGA  480
GCTCCTGGAC  CTAGGGCCCG  GCGTCGTCGC  CTCCTGGGCG  TCGCGGCAGA  GGGGAGTGGC  540
CCGCGCGGAA  AGCGCCGCGG  GACAGTCAGT  GACGAGGCCC  GGGGGTCGCC  GGGGCCACGA  600
CTTCTCGGAG  ACCGTCCTGC  GCTCTCTGGA  GACGCGCTGT  CCGCGCCCAG  GGTGGTGCCA  660
TGTGGGGCGC  TCGCCGCTCG  TCGTCTCCT  CATCCTGGAA  CGCCGCTTCG  CTCCTGCAGC  720
TGCTGCTGGC  TGCGCTGCTG  GCGGCGGGGG  CGAGGGCCCA  GCGGCGAGTA  CTGCCCACGGC  780
```

```
TGGCTGGACG CGCAGGGCGT CTGGCGCATC GGCTTCCAGT GTCCCGAGCG CTTCGACGGC   840
GGCGACGCCA CCATCTGCTG CGGCAGCTGC GCGTTGCGCT ACTGCTGCTC CAGCGCCGAG   900
GCGCGCCTGG ACCAGGGCGG CTGCGACAAT GACCGCCAGC AGGGCGCTGG CGAGCCTGGC   960
CGGGCGGACA AAGACGGGCC CCGACGGCTC GGCAGGGCTT CATGTCTTAG GGGTACCCAA   1020
GGAGACGGCG AGGGTGCGCC CCCACCCGTG AGGGCCTGGC AGCGGTGCTC CCCTGAAGGC   1080
TCCCCGAAAG GAAGGCAGCT CCTCAGGGCT TTCCCGGGGC TGCTGCCCCG TGCCAGACGC   1140
CGCGGATTCC CATCTTCTCC ACGCGGCGGC CCCTCTCCCC CGCCTTGCCC TGCCTTGCCC   1200
ATCTACGTGC CGTTCCTCAT TGTTGGCTCC GTGTTTGTCG CCTTTATCAT CTTGGGGTCC   1260
CTGGTGGCAG CCTGTTGCTG CAGATGTCTC CGGCCTAAGC AGGATCCCCA GCAGAGCCGA   1320
GCCCCAGGGG GTAACCGCTT GATGGAGACC ATCCCCATGA TCCCCAGTGC CAGCACCTCC   1380
CGGGGGTCGT CCTCACGCCA GTCCAGCACA GCTGCCAGTT CCAGCTCCAG CGCCAACTCC   1440
GGGGCCCGGG CGCCCCCAAC AAGGTCACAG ACCAACTGTT GCTTGCCGGA AGGGACCATG   1500
AACAACGTGT ATGTCAACAT GCCCACGAAT TTCTCTGTGC TGAACTGTCA GCAGGCCACC   1560
CAGATTGTGC CACATCAAGG GCAGTATCTG CATCCCCCAT ACGTGGGGTA CACGGTGCAG   1620
CACGACTCTG TGCCCATGAC AGCTGTGCCA CCTTTCATGG ACGGCCTGCA GCCTGGCTAC   1680
AGGCAGATTC AGTCCCCCTT CCCTCACACC AACAGTGAAC AGAAGATGTA CCCAGCGGTG   1740
ACTGTATAAC CGAGAGTCAC TGGTGGGTTC CTTTACTGAA GGGGACGAA GGCAGGGGTG   1800
GATTCTCGAG GTGGAAGTCC GCACATGTCG GTGGTATTTA TGGCACGATT CCTTTGGATG   1860
GCTTCATTTG CCCCCAGACT GTATGAAAAC ATCTCCGAAT TAGCATTTCT GGATATGTTT   1920
CATCCAGGGT ATCATTGATT TATGATGGAA AACCGGCCTC AGCTGGAGAT GACTGTGATG   1980
TTGCTGATGG GTGTATAACA AATGCTTGAG TCCGAAGTGC CCTTGAGATA TGGTTGACGA   2040
AAGAATTTTA TAAACTGATA AATTAAGGAT TTTTATTATG TTGTTATTAT TATTTCTTTT   2100
TTGTTGTTGA CTGCACAGGA TCAAAATGCC TGTTATCTCC CTTTTACTGG GACTTTTTTT   2160
TTTTTTTTTT TTTTTTTTAA TCAGACAGGG TCTTGCTCTG TTGCCCAGGC TGGAGTGCAG   2220
TGGTGCGATC TCGGCTCACT GCAACTTCAG CCTCCTGGAT TCAGGCAACA CTCCTGCCTC   2280
AGCCTCCCAC GTGGCTGGGA TTACAGGTGC CTGCCCCCAT GGCTAATTTT TTGTATTTTT   2340
TGTAGAGATG GGGTTTCACC ATGTTGGCTG GGCTGGTCTC ACTCTCCTGA CCTCAAGCAA   2400
TCTGCCTGTC TCAGCCTCCC AAAGTGCTGG GATTACAGGC GTGAGCCACC GCCCCCAGCC   2460
TGAGCCTTTT TTTTTTTCTA ATGCATCCAA GGTTAAGGGG AAGACGCAAA TAACAGGACT   2520
ATTCTAAAAG GAAACCTGTT TGAACTCTGT GAGATCAGTC ATCAGTCTCA GTATTCCACA   2580
GGCACACCTT AATTTCATTG TAAAAAGATA TATATATTTT GTCTATTTTT GTGCTTTTGG   2640
GGGCCTATTT TGTGCTTTTT TACCTTATGT AGAGATCTTA TTACAAAGTG ATTTTCTACA   2700
TTAAAAAGAG ACTGAAATAA ATTGTATAGT TACTTAACTA ATGAAGACAT TTCAGAACTC   2760
TGGGATGATT TTAATCTTGA AGTAGTAGGT GGTATAGTCA TAAAACCATT CATCCCCTTC   2820
TTGATTGTAT CTTAATTTTC TGGCTTTAAG GTGACATCTG AGAGGTAATG CATTCTTTTT   2880
TATATTGAAA TCATAAACTA TCACCCGCTG CTTCTCTGAG TTACTTTTAA TTTTGCCTTG   2940
TGGTTATGGT TTGGCGTTTC CTTCTGTTTG GTTTTCAGAG CCCCATGTCT ATATAGTCCT   3000
GAGTGCAAGT AATTACTATA CTTGTAAATG AAGATCAGTA TTTCTGCCTA GATCTGATAA   3060
AAAAATTTTC TTGTCTTAGT TATAAAAAAT CAAAGAAGTA TGTTACAAAG ATACTTAGTA   3120
TAGCTCCTCA GCCATAACCT GAGACTTGGG ATGAAATTTA AACCAGATAC GATTTACTTT   3180
GCAGATCATA AGGCTTTTTA TACTCTTGTT ATCAAAATGG CTTATTTTTC AGGCACTAAG   3240
GATTGTTAAG AGAAAAGCTT TTCAACGAAG GATTGCCTTT CTTCTCCCAC ACTGTTCTTG   3300
ATTTCCTCTC TCTTTCAGGC CTCAACAGGC ACTGTATTCA TTGCCAATGT TCCAAATTAT   3360
CAAATTCAAG TGAATTTATT TGTGTGTTCT TTACTTATAT AAAAAAAGAT AACTTTAAGG   3420
ATGTGCAAGT ACATTTCCAA CTGCTAGCAC AACCAGTATT TTGTAATTAA ACAAATCGGT   3480
GTATGGTATG GTCTTCTACA CATTTATGTC TATAGATATC TATCGATCAT CTTTCTATTC   3540
TGTTTCATGA CTGAATAATG TAAAACCAGT GTTGGCAATT GGTATCATCA ATGATACTCA   3600
TTTTTTAATA ACCAAAGGCA GGGGAAAATC ATTTTACTTA TTAATAAATA TTTTATGATG   3660
TGAAAAAAAA AAAAAAAAAA AAAAAAAAAA
```

A175 Protein sequence
Gene name:　　　　　　　　　　　　ESTs
Unigene number:　　　　　　　　　　Hs.128899
Protein Accession #:　　　　　　　none found
Signal sequence:　　　　　　　　　1-11
Transmembrane domains:　　　　　 402-424
Cellular Localization:　　　　　 not determined

```
1          11         21         31         41         51
|          |          |          |          |          |
MLSGFLMSPS TQHRAQYTPG GKKLFWEASI GAHTSRGRGS DRERESRPEA AGLLWDRAAA   60
GEAEKGNRGE PPAWIRAQQQ PRPPPAGQAP GTAAGGAQDP RLRPGRSRGR VRLPVKPPEA   120
SGRQPRGPSD CIPRFPSASA THKAVPKGTG PPAEDGDGLG APGPRABRRR LLGVAAEGSG   180
PRGKRRGTVS DEARGSPGPR LLGURPALSG DALSAPRVVP CGALAARPSP HPGTFLRSCS   240
CCWLRCWRRG RGPSGEYCHG WLDAQGVWRI GFQCPERFDG GDATICCGSC ALRYCCSSAE   300
ARLDQGGCDN DRQQGAGEPG RADKDGPRRL GRASCLRGTQ GDGEGAPPFV RAWQRCSPEG   360
SPRGRQLLRA FPGLLPRARR RGFPSSPRGG PSPLQRPALP IYVPFLIVGS VFVAFIILGS   420
LVAACCCRCL RPKQDPQQSR APGGNRLMET IPMIPSASTS RGSSSRQSST AASSSSSANS   480
GARAPPTRSQ TNCCLPEGTM NNVYVNMPTN FSVLNCQQAT QIVPEQGQYL HPPYVGYTVQ   540
HDSVPMTAVP PFMDGLQPGY RQIQSPFPHT NSEQKMYPAV TV
```

A176 DNA SEQUENCE
Gene name:　　ESTs, Weakly similar to CGHU7L collagen alpha 1(III) chain precursor [H.sapiens]
Unigene number:　　　　　　　　　Hs.19322
Probeset Accession #:　　　　　　AA088458
Nucleic Acid Accession #:　　　 AA088458
Coding sequence:　　　　　　　　　862-1995 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
GCCCTTGGAC ACTGACATGG ACTGAAGGAG TAGAATGGAG CACGAGGACA CTGACATGGA   60
```

```
      CTGAAGAAAA AGGAGCTGGA GCAGGAGAAG GAGGTGCTGC TGCAGGGGTTT GGAGATGATG   120
      GCGCGGGGCC GCGACTGGTA CCAGCAGCAG CTGCAACGAG TGCAGGAGCG CCAGCGCCGC   180
      CTGGGCCAGA GCAGAGCCAG CGCCGACTTT GGGGCTGCAG GGAGCCCCCG CCCACTGGGG   240
5     CGGCTACTGC CCAAGGTACA AGAGGTGGCC CGGTGCCTGG GGGAGCTGCT GGCTGCAGCC   300
      TGTGCCAGCC GGGCCCTGCC CCCGTCCTCC TCCGGGCCCC CCTGCCCTGC CCTGACGTCC   360
      ACCTCACCCC CGGTCTGGCA GCAGCAGACC ATCCTCATGC TGAAGGAGCA GAACCGACTC   420
      CTCACCCAGG AGGTGACCGA GAAGAGTGAG CGCATCACGC AGCTGGAGCA GGAGAAGTCG   480
      GCGCTCATTA AGCAGCTGTT TGAGGCCGGC GCCCTGAGCC AGCAGGACGG GGGACCTCTG   540
10    GATTCCACCT TCATCTAGTC CTTGTGGGCC GCGTGGGCCC CCAGGGCCAG CCTGGCACTC   600
      AGCCCTTCGA GGGTGGGCGC CCCATCGCAC CCACCCTCTC TGGCTGGAGA CCCCCGGCGG   660
      GCCCAGGCAC AGTCCCGGAG TGGGCGCCTT CCTGCCGCCG TTGCCAGATG GGCTCCCCAG   720
      GCCTGCCCCC GGCTGGTCCC CGCACCGAGC GCTTGACTCC GTTTKGGCTC CTGGTTGYTG   780
      ACATGGGCTG GGGGCTCTCT TGAGTCCGCA TAGTCCGCAG CTACTACTGG CCGCTGTCAG   840
15    TGGACAGTGG GGTACCCCTC CATGAGTTAG CGTCCCCCCG TTTCCAGCGG TGCCGCCCTG   900
      GGTCCCATCT TCAGGGAAAG GCACTGCCCA CGCCAGGCTG CACTTCCAAC AACGGGCAGC   960
      AGAGGGCGCG GGGCGGCTCC GACGCGGGTC CAAGGGCAGC TTCCCGCTCA ACCAGGGCAC  1020
      CAGGACGGAGG TGGCTGTAGC TCGGACGGGA GGAAGTAGAT GGAGGGGGTG GGGACGGCCT  1080
      GTAAGCGGGG GGTGCCTGCC TGGCTGGGGA GCCCCAGGGA TAGCGGTCGG ACTTCAGGTT  1140
20    CTGGCCAAGG CTGAGGGACC CTGGCTGCAG CGGATCGGCA CGCCGGGTGG GCGAGAGCTT  1200
      GGCCTGCATG TGCCTCCCAC AGACCCTGGG GTGATGGCCT TCCCCCTCTT GGCCGGGACG  1260
      TTGCCCCACG TTGAGTCCCA CACAACATCC TGTGAGCCTG GCTCCCCAGG AGGGCCCCCA  1320
      GACAGCTCCC AGGCACGTCA TAGGCAAAGC CTGTTTCCCC CGACTCAGGA TTTCCAAGGC  1380
      CTGGGGTCCT GCTCACCCCC CTTTGCTCTC ACGCCCAGCC TGTCCCCAGG TTTCAGCTGG  1440
25    GAGAGGCCAC CTCCCTCAGC CAAGGAAAAC GAGAACCCCC AGGGTACAGG AGGAGGCTGG  1500
      GGCAGGTCCC CTTGGGTGTC ACTCCCTCAG CCCCTGCCCA GGCCCACTCC CGCTGGTGCT  1560
      GGAGTACGCA CTGGTGGGGG GGCCCTGCTC AGCCCACCCT GGAGGGTCCG AGTGTCACCA  1620
      GAACCAGGGG CACGGCAACA GCATCGATGG GTTCTGCAGC CCAGGGCCCC CGATGCGGGG  1680
      TCAGTGTGTG TGGGGCGCAG GGCCTCCGAT GCGGGGTCAG TGCGTGGGGG GCGCAGGGCC  1740
30    CCCGATGCGG GGTCAGTGCG TGGGGGGCGC AGGGCCCCCT CGTGTCCAGG GCACTTTGGT  1800
      ACACTGTCCC ACAAGGCACC TGTCTCAGAG GAGGGGCCCT GGCAGGCAGC GTGGCAACTC  1860
      CCTTCCGGAG CCCAGCTCCA TGCTAACCTG CCCACAGCAA CCCCACAGAG CCACATTCCC  1920
      TGCTGCACCT GGTCTGCAGG GGTGTCCCAG GACAGGCCCA AGTCAGCCCA GCATGCAGCT  1980
      GCCCTCCTAC CCTGAAGATG GGAGTGGGCT TTCCAGGGGA CATAAGGATG TCAGGCCTGG  2040
35    ACCTCCTGGG CAGGAAAGGG TGCAGGTCCT GAGGGCCTGT GCCCCACAGC CCCAGCACCC  2100
      AGGTGGACTG CAGCGCAGTG GGTGGGCCAG TGGCAGCCAG GGAGAAGCCC CCCGTCAGCA  2160
      GGCTGGGGTC TGCCCACCAG GGCCTCCCCA CGTCTGCCTT TGAGGGTGCC TGCCATGCCC  2220
      TGGGGGGATCC TGGCATCCTT ACTGGACTGG AAGCAGGAGA CAGAACAGTG TCTGTCCCGG  2280
      GGTGACTTCA TCAGGAGACC GCCCACATAG AGCTGGACCC CGCAGCTGAA GCGGGAAATGT  2340
40    GAGACAGGCT GGCACCTCCG GAAAAACTGC CTTTCAGCCT TGGTGTTCCG TGCAAGGTGA  2400
      AAAGAAATAG GTCCTCCCAG TTTACAGCTT GAAATCAGGC TAGTGAGTGG CCCTGGAGAC  2460
      CACGAGGGGA GAATTTAAAG GCCCCGGCTG GCAGGGTCTA GGTGGCTGGC AGAGGCACAT  2520
      GCAGACCCTG CCTGGAGCCT GCCCTAGGAC GCTGGGCGGG TCAGTCTCCG TGCAGGATGT  2580
      GAGCAGCGTC CCTGGGGCTCT ATCCGCGAGG TGCCAGTAGC GTGTGCAGGT ACATACACGT  2640
45    GCGTGCACAC TGTGATGACA CCCGGGAAATG TCTCAGGATG TTGAAATGTG TCCTTGGGGG  2700
      CAGAAGTGTC CCCAGTTGAG AATCTGCCCC AGAGGAACAC ACCCACACCA GGCCTCAGGA  2760
      TTTTGTGTTG ATCAAGTTCC AAGGAAAAGG AACATCTCAG CCGGGCGTGG TGGTTCACGC  2820
      CTGGAATCCC AGCACTTGAG GCCAGGAGTT CCAGAGCAGC CTGGGCAACG CAGTGAGAGA  2880
50    CCCCATCTCT ACAARAAAAA AAAAAGAAAG AAAGAAAATG AGAGATCCAG GTTTAAAAAT  2940
      TCATAAACAC CACAAGGAAA CAATACACTA TGAGACCCAG CAGAAGCAAC AGATTGACTC  3000
      TAGACCCAGA TACTAGAATT ATCAGAGAGA ATATAAAGTA ACAGTGTTTT ATATATCTAA  3060
      AGAAATAAAA GAGATTTCTG GAAACATGAA AAAAAA
```

A177 DNA SEQUENCE

Gene name:	Cadherin 3, P-cadherin (placental)
Unigene number:	Hs.2877
Probeset Accession #:	X63629
Nucleic Acid Accession #:	X63629
Coding sequence:	54-2543 (start and stop codons are underlined)

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      GCGGAACACC GGCCCGCCGT CGCGGCAGCT GCTTCACCCC TCTCTCTGCA GCCATGGGGC   60
      TCCCTCGTGG ACCTCTCGCG TCTCTCCTCC TTCTCCAGGT TTGCTGGCTG CAGTGCGCGG  120
      CCTCCGAGCG GTGCCGGGCG GTCTTCAGGG AGGCTGAAGT GAACCTTGGAG GCGGGAGGCG  180
      CGGAGCAGGA GCCCGGCCAG GCGCTGGGGA AAGTATTCAT GGGCTGCCCT GGGCAAGAGC  240
      CAGCTCTGTT TAGCACTGAT AATGATGACT TCACTGTGCG GAATGGCGAG ACAGTCCAGG  300
      AAAGAAGGTC ACTGAAGGAA AGGAATCCAT TGAAGATCTT CCCATCCAAA CGTATCTTAC  360
      GAAGACACAA GAGAGATTGG GTGGTTGCTC CAATATCTGT CCCTGAAAAT GGCAAGGGTC  420
      CCTTCCCCCA GAGACTGAAT CAGCTCAAGT CTAATAAAGA TAGAGACACC AAGATTTTCT  480
      ACAGCATCAC GGGGCCGGGG GCAGACAGCC CCCCTGAGGG TGTCTTCGCT GTAGAGAAGG  540
      AGACAGGCTG GTTGTTGTTG AATAAGCCAC TGGACCGGGA GGAGATTGCC AAGTATGAGC  600
      TCTTTGGCCA CGCTGTGTCA GAGAATGGTG CCTCAGTGGA GGACCCCATG AACATCTCCA  660
      TCATCGTGAC CGACCAGAAT GACCACAAGC CCAAGTTTAC CCAGGACACC TTCCGAGGGA  720
      GTGTCTTAGA GGGAGTCCTA CCAGGTACTT CTGTGATGCA GGTGACAGCC ACAGATGAGG  780
75    ATGATGCCAT CTACACCTAC AATGGGGTGG TTGCTTACTC CATCCATAGC CAAGAACCAA  840
      AGGACCCACA CGACCTCATG TTCACAATTC ACCGGAGCAC AGGCACCATC AGCGTCATCT  900
      CCAGTGGCCT GGACCGGGAA AAAGTCCCTG AGTACACACT GACCATCCAG GCCACAGACA  960
      TGGATGGGGA CGGCTCCACC ACCACGGCAG TGGCAGTAGT GGAGATCCTT GATGCCAATG  1020
      ACAATGCTCC CATGTTTGAC CCCCAGAAGT ACGAGGCCCA TGTGCCTGAG AATGCAGTGG  1080
80    GCCATGAGGT GCAGAGGCTG ACGGTCACTG ATCTGGACGC CCCCAACTCA CCAGCGTGGC  1140
      GTGCCACCTA CCTTATCATG GGCGGTGACG ACGGGGACCA TTTTACCATC ACCACCCACC  1200
      CTGAGAGCAA CCAGGGCATC CTGACAACCA GGAAGGGTTT GGATTTTGAG GCCAAAAACC  1260
      AGCACACCCT GTACGTTGAA GTGACCAACG AGGCCCCTTT TGTGCTGAAG CTCCCAACCT  1320
      CCACAGCCAC CATAGTGGTC CACGTGGAGG ATGTGAATGA GGCACCTGTG TTTGTCCCAC  1380
```

```
        CCTCCAAAGT CGTTGAGGTC CAGGAGGGCA TCCCCACTGG GGAGCCTGTG TGTGTCTACA 1440
        CTGCAGAAGA CCCTGACAAG GAGAATCAAA AGATCAGCTA CCGCATCCTG AGAGACCCAG 1500
        CAGGGTGGCT AGCCATGGAC CCAGACAGTG GGCAGGTCAC AGCTGTGGGC ACCCTCGACC 1560
   5    GTGAGGATGA GCAGTTTGTG AGGAACAACA TCTATGAAGT CATGGTCTTG GCCATGGACA 1620
        ATGGAAGCCC TCCCACCACT GGCACGGGAA CCCTTCTGCT AACACTGATT GATGTCAACG 1680
        ACCATGGCCC AGTCCCTGAG CCCCGTCAGA TCACCATCTG CAACCAAAGC CCTGTGCGCC 1740
        ACGTGCTGAA CATCACGGAC AAGGACCTGT CTCCCCACAC CTCCCCTTTC CAGGCCCAGC 1800
        TCACAGATGA CTCAGACATC TACTGGACGG CAGAGGTCAA CGAGGAAGGT GACACAGTGG 1860
   10   TCTTGTCCCT GAAGAAGTTC CTGAAGCAGG ATACATATGA CGTGCACCTT TCTCTGTCTG 1920
        ACCATGGCAA CAAAGAGCAG CTGACGGTGA TCAGGGCCAC TGTGTGCGAC TGCCATGGCC 1980
        ATGTCGAAAC CTGCCCTGGA CCCTGGAAAG GAGGTTTCAT CCTCCCTGTG CTGGGGGCTG 2040
        TCCTGGCTCT GCTGTTCCTC CTGCTGGTGC TGCTTTTGTT GGTGAGAAAG AAGCGGAAGA 2100
        TCAAGGAGCC CCTCCTACTC CCAGAAGATG ACACCCGTGA CAACGTCTTC TACTATGGCG 2160
   15   AAGAGGGGGG TGGCGAAGAG GACCAGGACT ATGACATCAC CCAGTCCCTG CGAGGTCTGG 2220
        AGGCCAGGCC GGAGGTGGTT CTCCGCAATG ACGTGGCACC AACCATCATC CCGACACCCA 2280
        TGTACCGTCC TAGGCCAGCC AACCCAGATG AAATCGGCAA CTTTATAATT GAGAACCTGA 2340
        AGGCGGCTAA CACAGACCCC ACAGCCCCGC CCTACGACAC CCTCTTGGTG TTCGACTATG 2400
        AGGGCAGCGG CTCCGACGCC GCGTCCCTGA GCTCCCTCAC CTCCTCCGCC TCCGACCAAG 2460
   20   ACCAAGATTA CGATTATCTG AACGAGTGGG GCAGCCGCTT CAAGAAGCTG GCAGACATGT 2520
        ACGGTGGCGG GGAGGACGAC TAGGCCGGCCT GCCTGCAGGG CTGGGGACCA AACGTCAGGC 2580
        CACAGAGCAT CTCCAAGGGG TCTCAGTTCC CCCTTCAGCT GAGGACTTCG GAGCTTGTCA 2640
        GGAAGTGGCC GTAGCAACTT GGCGGAGACA GGCTATGAGT CTGACGTTAG AGTGGTTGCT 2700
        TCCTTAGCCT TTCAGGATGG AGGAATGTGG GCAGTTTGAC TTCAGCACTG AAAACCTCTC 2760
   25   CACCTGGGCC AGGGTTGCCT CAGAGGCCAA GTTTCCAGAA GCCTCTTACC TGCCGTAAAA 2820
        TGCTCAACCC TGTGTCCTGG GCCTGGGCCT GCTGTGACTG ACCTACAGTG GACTTTCTCT 2880
        CTGGAATGCA ACCTTCTTAG GCCTCCTGGT GCAACTTAAT TTTTTTTTTT AATGCTATCT 2940
        TCAAAACGTT AGAGAAAGTT CTTCAAAAGT GCAGCCCAGA GCTGCTGGGC CCACTGGCCG 3000
        TCCTGCATTT CTGGTTTCCA GACCCCAATG CCTCCCATTC GGATGGATCT CTGCGTTTTT 3060
   30   ATACTGAGTG TGCCTAGGTT GCCCCTTATT TTTTATTTTC CCTGTTGCGT TGCTATAGAT 3120
        GAAGGGTGAG GACAATCGTG TATATGTACT AGAACTTTTT TATTAAAGAA A
```

<u>A178 Protein sequence:</u>

Gene name:	Cadherin 3, P-cadherin (placental)
Unigene number:	Hs.2877
Probeset Accession #:	X63629
Protein Accession #:	CAA45177
Signal sequence:	1-24
Transmembrane domain:	659-675

Cellular localization: plasma membrane

```
        1         11        21        31        41        51
        |         |         |         |         |         |
        MGLPRGPLAS LLLLQVCWLQ CAASEPCRAV PREAEVTLEA GGAEQEPGQA LGKVFMGCPG  60
        QEPALPSTDN DDFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG 120
        KGPFPQRLNQ LKSNKDRDTK IFYSITGPGA DSPPEGVFAV EKETGWLLLN KPLDREEIAK 180
        YELFGHAVSE NGASVEDPMN ISIIVTDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT 240
        DEDDAIYTYN GVVAYSIHSQ EPKDPHDLMF TIHRSTGTIS VISSGLDREK VPEYTLTIQA 300
        TDMDGDGSTT TAVAVVHILD ANDNAPMFDP QKYEAHVPEN AVGHEVQRLT VTDLDAPNSP 360
        AWRATYLIMG GDDGDHFTIT THPESNQGIL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL 420
        PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGIPTGEPVC VYTAEDPDKE NQKISYRILR 480
        DPAGWLAMDP DSGQVTAVGT LDREDEQFVR NNIYEVMVLA MDNGSPPTTG TGTLLLTLID 540
        VNDHGPVPEP RQITICNQSP VRHVLNITDK DLSPHTSPPQ AQLTDDSDIY WTAEVNEEGD 600
        TVVLSLKKPL KQDTYDVHLS LSDHENKEQL TVIRATVCDC HGHVETCPGP WKGGFILPVL 660
        GAVLALLFLL LVLLLLVRKK RKIKEPLLLP EDDTRDNVFY YGEEGGGEED QDYDITQLER 720
        GLEARPEVVL RNDVAPTIIP TPMYRPRPAN PDEIGNFIIE NLKAANTDPT APPYDTLLWF 780
        DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD
```

<u>A179 DNA SEQUENCE</u>

Gene name:	endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 7
Unigene number:	Hs.258583
Probeset Accession #:	NM_012152
Nucleic Acid Accession #:	NM_012152
Coding sequence:	43-1104 (underlined sequences correspond to start and stop codons)

```
        1         11        21        31        41        51
        |         |         |         |         |         |
        CTTCTTTAAA TTTCTTTCTA GGATGTTCAC TTCTTCTCCA CAATGAATGA GTGTCACTAT  60
        GACAAGCACA TGGACTTTTT TTATAATAGG AGCAACACTG ATACTGTCGA TGACTGGACA 120
        GGAACAAAGC TTGTGATTGT TTTGTGTGTT GGGACGTTTT TCTGCCTGTT TATTTTTTTT 180
        TCTAATTCTC TGGTCATCGC GGCAGTGATC AAAAACAGAA AATTTCATTT CCCCTTCTAC 240
        TACCTGTTGG CTAATTTAGC TGCTGCCGAT TTCTTCGCTG GAATTGCCTA TGTATTCCTG 300
        ATGTTTAACA CAGGCCCAGT TTCAAAAACT TTGACTGTCA ACCGCTGGTT TCTCCGTCAG 360
        GGGCTTCTGG ACAGTAGCTT GACTGCTTCC CTCACCAACT TGCTGGTTAT CGCCGTGGAG 420
        AGGCACATGT CAATCATGAG GATGCGGGTC CATAGCAACC TGACCAAAAA GAGGGTGACA 480
        CTGCTCATTT TGCTTGTCTG GGGCATCGCC ATTTTTATGG GGCGGTCCC CACACTGGGC 540
        TGGAATTGCC TCTGCAACAT CTCTGCCTGC TCTTCCCTGG CCCCCATTTA CAGCAGGAGT 600
        TACCTTGTTT TCTGGACAGT GTCCAACCTC ATGGCCTTCC TCATCATGGT TGTGGTGTAC 660
        CTGCGGATCT ACGTGTACGT CAAGAGGAAA ACCAACGTCT TGTCTCCGCA TACAAGTGGG 720
        TCCATCAGCC GCCGGAGGAC ACCCATGAAG CTAATGAAGA CGGTGATGAC TGTCTTAGGG 780
```
```
                                917
```

```
GCGTTTGTGG TATGCTGGAC CCCGGGCCTG GTGGTTCTGC TCCTCGACGG CCTGAACTGC   840
AGGCAGTGTG GCGTGCAGCA TGTGAAAAGG TGGTTCCTGC TGCTGGCGCT GCTCAACTCC   900
GTCGTGAACC CCATCATCTA CTCCTACAAG GACGAGGACA TGTATGGCAC CATGAAGAAG   960
ATGATCTGCT GCTTCTCTCA GGAGAACCCA GAGAGGCGTC CCTCTCGCAT CCCCTCCACA  1020
GTCCTCAGCA GGAGTGACAC AGGCAGCCAG TACATAGAGG ATAGTATTAG CCAAGGTGCA  1080
GTCTGCAATA AAAGCACTTC CTAAACTCTG GATGCCTCTC GGCCCACCCA GGTGATGACT  1140
GTCTTAGG
```

<u>A180 Protein sequence:</u>
Gene name: endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 7
Unigene number: Hs.258583
Probeset Accession #:   NM_012152
Protein Accession #:   NP_036284
Signal sequence: none found
Transmembrane domains: 31-53, 66-88, 150-172, 190-211, 239-261, 277-295
Cellular Localization: plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
MNECHYDKHM DFFYNRSNTD TVDDWTGTXL VIVLCVGTFF CLFIFFSNSL VIAAVIKNRK   60
FHFPFYYLIA NLAAADPFAG IAYVFLMFNT GPVSKTLTVN RWFLRQGLLD SSLTASLTNL  120
LVIAVERHMS IMRMRVHSNL TKKRVTLLIL LVWAIAIFMG AVPTLGWNCL CNISACSSLA  180
PIYSRSYLVF WTVSNLMAFL IMVVVYLRIY VYVRKTNVL SPETSGSISR RRTPMKLMKT  240
VMTVLGAPVV CWTPGLVVLL LDGLNCRQCG VQHVKRWFLL LALLNSVVNP IIYSYKDEDM  300
YGTMKKMICC PSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK STS
```

<u>A181 DNA SEQUENCE</u>
Gene name: ESTs
Unigene number: Hs.162859
Probeset Accession #:   AA569531
Nucleic Acid Accession #: AA569533
Coding sequence: 1-504 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGACCTACA GTTACTCATT TTTCAGGCCT GAGTTGATCG TTAATCATCT TAATTATGTT   60
CATTCTGAAG CCAACAGGAG AACCAAGACC AAAACTTTAT TGTCTCTGCT TTCATTTCTT  120
GATGAAACCT CTGGACTAAG CACACATCTT CCTTGTTTAT CTCTCTCAAA GGAGTGTGGA  180
GTGCTTCATC TGGACATCCA CGGGAAGAAG GAAGACATGA GAATCACCCA ACAGTCTTCC  240
CAGCTATACC TGTGGGACAT GGGTGGTTTT ACAATATTTA AGAACCTGTG GATGAGCCTC  300
ATACCCAGAG GGAACAAACG CTCCCCAAAA AGAGTTACAG AAAACCATCCT GAGAGATTTT  360
AAGCAGAAGC AAAGTTCAAA GATCCAAGAG GAGAGACGAA GAGAGTCTGC AGGACCAAAC  420
CTCTCTTCAT TCTGGTTTGT GGGGAATGCT GGAAGAGGCA ACAGGCCCCA GATTTGGGCA  480
GGAAGTAAAC AGTTTTCAGG CTGAGGCCAA TCTGAGCAGG AACATTCCAA TATTTCTTCA  540
GCTACGTTGT CCCAGCACTT CACTGGTTAA CCTTTTATGT CCACCATTTG TGGATTTCAC  600
AGCTACTTGT CAATGGTGAA TATTGATCAT CATCATTATC TGCTGAGCTG CTACCATATC  660
CCAGCTACTC CTTGCATGTT GTTCATTATT TTCTCAACAC TCAGCATATT TGCAATATGT  720
TATGTAATAT CACAGACAAG GAAACTGAAC GCAGAAATGT TTTATTTCTT GCCAAACATC  780
ACATGAGGAT GAACAATGAA ACCGATTTGA AACCAGGATT GTCTGATTCC AACATCTCTG  840
GGTCCTTTTT CACTCTGATA TGCTGCAATT AAAAAGCCAT TTCTAAGACT GT
```

<u>A182 Protein sequence:</u>
Gene name: ESTs
Unigene number: Hs.162859
Probeset Accession #:   AA569531
Protein Accession #:   none found
Signal sequence: 1-46
Transmembrane domains: none found
Cellular Localization: not determined

```
1          11         21         31         41         51
|          |          |          |          |          |
MTYSYSFFRP ELIVNHLNYV HSEANRRTKT KTLLSLLSFL DETSGLSTHL PCLSLSKECG   60
VLHLDIHGKK EDMRITQQSS QLYLWDMGGF TIFKNLWMSL IPRGNKRSPK RVTETILRDF  120
KQKQSSKIQE ERRRESAGPN LSSFWFVGNA GRGDRPQIWA GSKQFSG
```

<u>A183 DNA SEQUENCE</u>
Gene name: ESTs
Unigene number: Hs.179809
Probeset Accession #:   N95796
Nucleic Acid Accession #: XM_050197
Coding sequence: 310-1971 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
TCACACGTGC CAAGGGGCTG GCTCAGCGGA ACCAGCCTGC ACGCGCTGGC TCCGGGTGAC   60
AGCCGCGCGC CTCGGCCAGG ATCTGAGTGA TGAGACGTGT CCCCACTGAG GTGCCCCACA  120
GCAGCAGGTG TTGAGCATGG GCTGAGAAGC TGGACCGGCA CCAAAGGGCT GGCAGAAATG  180
GGCGCCTGGC TGATTCCTAG GCAGTTGGCG GCAGCAAGGA GGAGAGGCCG CAGCTTCTGG  240
AGCAGAGCCG AGACGAAGCA GTTCTGGAGT GCCTGAACGG CCCCCTGAGC CCTACCCGCC  300
```

918

```
TGGCCCACTA TGGTCCAGAG GCTGTGGGTG AGCCGCCTGC TGCGGCACCG GAAAGCCCAG   360
CTCTTGCTGG TCAACCTGCT AACCTTTGGC CTGGAGGTGT GTTTGGCCGC AGGCATCACC   420
TATGTGCCGC CTCTGCTGCT GGAAGTGGGG GTAGAGGAGA AGTTCATGAC CATGGTGCTG   480
GGCATTGGTC CAGTGCTGGG CCTGGTCTGT GTCCCGCTCC TAGGCTCAGC CAGTGACCAC   540
TGGCGTGGAC GCTATGGCCG CCGCCGGCCC TTCATCTGGG CACTGTCCTT GGGCATCCTG   600
CTGAGCCTCT TTCTCATCCC AAGGGCCGGC TGGCTAGCAG GGCTGCTGTG CCCGGATCCC   660
AGGCCCCTGG AGCTGGCACT GCTCATCCTG GGCGTGGGGC TGCTGGACTT CTGTGGCCAG   720
GTGTGCTTCA CTCCACTGGA GGCCCTGCTC TCTGACCTGT TCCGGGACCC GGACCACTGT   780
CGCCAGGCCT ACTCTGTCTA TGCCTTCATG ATCAGTCTTG GGGGCTGCCT GGGCTACCTC   840
CTGCCTGCCA TTGACTGGGA CACCAGTGCC CTGGCCCCGT ACCTGGGCAC CCAGGAGGAG   900
TGCCTCTTTG GCCTGCTCAC CCTCATCTTC CTCACCCTGCG TAGCAGCCAC ACTGCTGGTG   960
GCTGAGGAGG CAGCGCTGGG CCCCACCGAG CCAGCAGAAG GGCTGTCGGC CCCCTCCTTG  1020
TCGCCCCACT GCTGTCCATG CGGGGCCCGC TTGGCTTTCC GGAACCTGGG CGCCCTGCTT  1080
CCCCGGCTGC ACCAGCTGTG CTGCCGCATG CCCCGCACCC TGCGCCGGCT CTTCGTGGCT  1140
GAGCTGTGCA GCTGGATGGC ACTCATGACC TTCACGCTGT TTTACACGGA TTTCGTGGGC  1200
GAGGGGCTGT ACCAGGGCGT GCCCAGAGCT GAGCCGGGCA CCGAGGCCCG GAGACACTAT  1260
GATGAAGGCG TTCGGATGGG CAGCCTGGGG CTGTTCCTGC AGTGCGCCAT CTCCCTGGTC  1320
TTCTCTCTGG TCATGGACCG GCTGGTGCAG CGATTCGGCA CTCGAGCAGT CTATTTGGCC  1380
AGTGTGGCAG CTTTCCCTGT GGCTGCCGGT GCCACATGCC TGTCCCACAG TGTGGCCGTG  1440
GTGACAGCTT CAGCCGCCCT CACCGGGTTC ACCTTCTCAG CCCTGCCAGAT CCTGCCCTAC  1500
ACACTGGCCT CCCTCTACCA CCGGGAGAAG CAGGTGTTCC TGCCCAAATA CCGAGGGGAC  1560
ACTGGAGGTG CTAGCAGTGA GGACAGCCTG ATGACCAGCT TCCTGCCAGG CCCTAAGCCT  1620
GGAGCTCCCT TCCCTAATGG ACACGTGGGT GCTGGAGGCA GTGGCCTGCT CCCACCTCCA  1680
CCCGCGCTCT GCGGGGCCTC TGCCTGTGAT GTCTCCGTAC GTGTGGTGGT GGGTGAGCCC  1740
ACCGAGGCCA GGGTGGTTCC GGGCCGGGGC ATCTGCCTGG ACCTCGCCAT CCTGGATAGT  1800
GCCTTCCTGC TGTCCCAGGT GGCCCCATCC CTGTTTATGG GCTCCATTGT CCAGCTCAGC  1860
CAGTCTGTCA CTGCCTATAT GGTGTCTGCC GCAGGCCTGG GTCTGGTCGC CATTTACTTT  1920
GCTACACAGG TAGTATTTGA CAAGAGCGAC TTGGCCAAAT ACTCAGCGTA GAAAACTTCC  1980
AGCACATTGG GGTGGAGGGC CTGCCCTCACT GGGTCCCAGC TCCCCGCTCC TGTTAGCCCG  2040
ATGGGGCTGC CGGGCTGGCC GCCAGTTTCT GTTGCTGCCA AAGTAATGTG GCTCTCTGCT  2100
GCCACCCTGT GCTGCTGAGG TGCGTAGCTG CACAGCTGGG GGCTGGGGCG TCCCTCTCCT  2160
CTCTCACGGG TCTCTAGGGC TGCCTGACTG GAGGCCTTCC AAGGGGGTTT CAGTCTGGAC  2220
TTATACAGGG AGGCCAGAAG GGCTCCATGC ACTGGAATGC GGGGACTCTG CAGGTGGATT  2280
ACCCAGGCTC AGGGTTAACA GCTAGCCTCC TAGTTGAGAC ACACCTAGAG AAGGGTTTTT  2340
GGGAGCTGAA TAAACTCAGT CACCTGGTTT CCCATCTCTA AGCCCCTTAA CCTGCAGCTT  2400
CGTTTAATGT AGCTCTTGCA TGGGAGTTTC TAGGATGAAA CACTCCTCCA TGGGATTTGA  2460
ACATATGAAA GTTATTTGTA GGGGAAGAGT CCTGAGGGGC AACACACAAG AACCAGGTCC  2520
CCTCAGCCCC ACAGGCACTG GTCTTTTTTG CTNGANTCCA CCCCCCCCCT CTTTACCCTT  2580
TT
```

A184 Protein sequence:

Gene name:	ESTs
Unigene number:	Hs_179809
Probeset Accession #:	N95796
Protein Accession #:	XP_050197
Signal sequence:	none
Transmembrane domains:	21-43, 53-75, 90-112, 125-147, 160-182, 199-221, 323-345, 350-372, 379-401
Cellular Localization:	plasma membrane

```
  1          11         21         31         41         51
  |          |          |          |          |          |
MVQRLWVSRL LRHRKAQLLL VNLLTFGLEV CLAAGITYVP PLLLEVGVEE KFMTMVLGIG    60
PVLGLVCVPL LGSASDHWRG RYGRRRPFIW ALSLGILLSL FLIPRAGWLA GLLCFDPRPL   120
ELALLILGVG LLDFCGQVCF TPLEALLSDL FRDPDHCRQA YSVYAFMISL GGCLGYLLPA   180
IDWDTSALAP YLGTQEECLF GLLTLIFLTC VAATLLVAEE AALGPTEPAE GLSAPSLSPH   240
CCPCRARLAF RNLGALLPRL HQLCCRMPRT LRRLFVAELC SWMALMTFTL FYTDFVGEGL   300
YQGVPRAEPG TEARRHYDEG VRMGSLGLFL QCAISLVFSL VMDRLVQRFG TRAVYLASVA   360
AFPVAAGATC LSHSVAVVTA SAALTGFTFS ALQILPYTLA SLYHREKQVF LPKYRGDTGG   420
ASSEDSLMTS FLPGPKPGAP FPNGHVGAGG SGLLPPPPAL CGASACDVSV RVVVGEPTEA   480
RVVPGRGICL DLAILDSAFL LSQVAPSLFM GSIVQLSQSV TAYMVSAAGL GLVAIYFATQ   540
VVFDKSDLAK YSA
```

A185 DNA SEQUENCE

Gene name:	ESTs
Unigene number:	Hs.11260
Probeset Accession #:	R73640
Nucleic Acid Accession #:	AK002126
Coding sequence:	1-1593 (underlined sequences correspond to start and stop codons)

```
  1          11         21         31         41         51
  |          |          |          |          |          |
ATGGTTCGCC GGGGGCTGCT TGCGTGGATT TCCCGGGTGG TGGTTTTGCT GGTGCTCCTC    60
TGCTGTGCTA TCTCTGTCCT GTACATGTTG GCCTGCACCC CAAAAGGTGA CGAGGAGCAG   120
CTGGCACTGC CCAGGGCCAA CAGCCCCACG GGGAAGGAGG GGTACCAGGC CGTCCTTCAG   180
GAGTGGGAGG AGCAGCACCG CAACTACGTG AGCAGCCTGA AGCGGCAGAT CGCACAGCTC   240
AAGGAGGAGC TGCAGGAGAG GAGTGAGCAG CTCAGGAATG GGCAGTACCA AGCCAGCGAT   300
GCTGCTGGCC TGGGTCTGGA CAGGAGCCCC CCAGAGAAAA CCCAGAGCCGA CCTCCTGGCC   360
TTCCTGCACT CGCAGGTGGA CAAGGCAGAG GTGAATGCTG GCGTCAAGCT GGCCACAGAG   420
TATGCAGCAG TGCCTTTCGA TAGCTTTACT CTACAGAAGG TGTACCAGCT GGAGACTGGC   480
CTTACCCGCC ACCCGAGGA GAAGCCTGTG AGGAAGGACA AGCGGGATCG GTTGGTGGAA   540
GCCATTGAAT CAGCCTTGGA GACCCTGAAC AATCCTGCAG AGAACAGACGA CAATCACCGT   600
```

```
CCTTACACGG CCTCTGATTT CATAGAAGGG ATCTACCGAA CAGAAAGGGA CAAAGGGACA   660
TTGTATGAGC TCACCTTCAA AGGGGACCAC AAACACGAAT TCAAACGGCT CATCTTATTT   720
CGACCATTCG GCCCCATCAT GAAAGTGAAA AATGAAAAGC TCAACATGGC CAACACGCTT   780
ATCAATGTTA TCGTGCCTCT AGCAAAAGGG GTGGACAAGT TCCGGCAGTT CATGCAGAAT   840
TTCAGGGAGA TGTGCATTGA GCAGGATGGG AGAGTCCATC TCACTGTTGT TTACTTTGGG   900
AAAGAAGAAA TAAATGAAGT CAAAGGAATA CTTGAAAACA CTTCCAAAGC TGCCAACTTC   960
AGGAACTTTA CCTTCATCCA GCTGAATGGA CAATTTTCTC GGGGAAAGGG ACTTGATGTT  1020
GGAGCCCGCT TCTGGAAGGG AAGCAACGTC CTTCTCTTTT TCTGTGATGT GGACATCTAC  1080
TTCACATCTG AATTCCTCAA TACGTGTAGG CTGAATACAC AGCCAGGGAA GAAGGTATTT  1140
TATCCAGTTC TTTTCAGTCA GTACAATCCT GGCATAAATG ACGGCCACCA TGATGCAGTC  1200
CCTCCCTTGG AACAGCAGCT GGTCATAAAG AAGGAAACTG GATTTTGGAG AGACTTTGGA  1260
TTTGGGATGA CGTGTCAGTA TCGGTCAGAC TTCATCAATA TAGGTGGGTT TGATCTGGAC  1320
ATCAAAGGCT GGGGCGGAGA GGATGTGCAC CTTTATCGGA AGTATCTCCA CAGCAACCTC  1380
ATAGTGGTAC GGACGCCTGT GCGAGGACTC TTCCACCTCT GGCATGAGAA GCGCTGCATG  1440
GACGAGCTGA CCCCCGAGCA GTACAAGATG TGCATGCAGT CCAAGGCCAT GAACGAGGCA  1500
TCCCACGGCC AGCTGGGCAT GCTGGTGTTC AGGCACGAGA TAGAGGCTCA CCTTCGCAAA  1560
CAGAAACAGA AGACAAGTAG CAAAAAAACA TGA                                1560
```

```
 1          11         21         31         41         51
 |          |          |          |          |          |
MVRRGLLAWI SRVVVLLVLL CCAISVLYML ACTPKGDEEQ LALPRANSPT GKEGYQAVLQ    60
EWEEQHRNYV SSLKRQIAQL KEELQERSEQ LRNGQYQASD AAGLGLDRSP PEKTQADLIA   120
PLHSQVDKAE VNAGVKLATS YAAVPFDSFT LQKVYQLETG LTRHPEEKPV RKDKRDELVE   180
AIESALETLN NPAENSPNHR PYTASDFIEG IYRTERDKGT LYELTFKGDH KHEFKRLILF   240
RPPGPIMKVK NEKLNMANTL INVIVPLAKR VDKFRQFMQN FREMCIEQDG RVHLTVVYFG   300
KEEINEVKGI LENTSKAANF RNFTFIQLNG EFSRGKGLDV GARFWKGSNV LLFFCDVDIY   360
FTSEFLNTCR LNTQPGKKVF YPVLFSQYNP GIIYGHHDAV PPLEQQLVIK KETGFWRDFG   420
FGMTCQYRSD PINIGGFDLD IKGWGGEDVH LYRKYLHSNL IVVRTPVRGL FHLWHEKRCM   480
DELTPEQYKM CMQSKAMNEA SHGQLGMLVF RHEIEAHLRK QKQKTSSKKT
```

```
 1          11         21         31         41         51
 |          |          |          |          |          |
ATGATTCCTG TATTGACATC AAAAAAAGCA AGTGAATTAC CAGTCAGTGA AGTTGCAAGC    60
ATTCTCCAAG CTGATCTTCA GAATGGTCTA AACAAATGTG AAGTTAGTCA TAGGCGAGCC   120
TTTCATGGCT GGAATGAGTT TGATATTAGT GAAGATGAGC CACTGTGGAA GAAGTATATT   180
TCTCAGTTTA AAAATCCCCT TATTATGCTG CTTCTGGCTT CTGCAGTCAT CAGTGTTTTA   240
ATGCATCAGT TTGATGATGC CGTCAGTATC ACTGTGGCAA TACTTATCGT TGTTACAGTT   300
GCCTTTGTTC AGGAATATGG TTCAGAAAAA TCTCTTGAAG AATTGAGTAA ACTTGTGCCA   360
CCAGAATGCC ATTGTGTGCG TGAAGGAAAA TTGGAGCATA CACTTGCCCG AGACTTGGTT   420
CCAGGTGATA CAGTTTGCCT TTCTGTTGGG GATAGAGTTC CTGCTGACTT ACGCTTGTTT   480
GAGGCTGTGG ATCTTTCCAT TGATGAGTCC AGCTTGACAG GTGAGACAAC GCCTTGTTCT   540
AAGGTGACAG CTCCTCAGCC AGCTGCAACT AATGGAGATC TTGCATCGAG AAGTAACATT   600
GCCTTTATGG GAACACTGGT CAGATGTGGC AAAGCAAAGG GTGTTGTCAT TGGAACAGGA   660
GAAAATTCTG AATTTGGGGA GGTTTTTAAA ATGATGCAAG CAGAAGAGGC ACCAAAAACC   720
CCTCTGCAGA AGAGCATGGA CCTCTTAGGA AAACAACTTT CCTTTTACTC CTTTGGTATA   780
ATAGGAATCA TCATGTTGGT TGGCTGGTTA CTGGGAAAAG ATATCCTGGA AATGTTTACT   840
ATTAGTGTAA GTTTGGCTGC AGCAGCAATT CCTGAAGGTC TCCCCATTGT GGTCACAGTG   900
ACGCTAGCTC TTGGTGTTAT GAGAATGGTG AAGAAAGGG CCATTGTGAA AAAGCTGCCT   960
ATTGTTGAAA CTCTGGGCTG CTGTAATGTG ATTTGTTCAG ATAAAACTGG AACACTGACG  1020
AAGAATGAAA TGACTGTTAC TCACATATTT ACTTCAGATG GTCTGCATGC TGAGGTTACT  1080
GGAGTTGGCT ATAATCAATT TGGGGAAGTG ATTGTTGATG GTGATGTTGT TCATGGATTC  1140
TATAACCCAG CTGTTAGCAG AATTGTTGAG GCGGGCTGTG TGTGCAATGA TGCTGTAATT  1200
AGAAACAATA CTCTAATGGG GAAGCCAACA GAAGGGGCCT TAATTGCTCT TGCAATGAAG  1260
ATGGGTCTTG ATGGACTTCA ACAAGACTAC ATCAGAAAAG CTGAATACCC TTTTAGCTCT  1320
GAGCAAAAGT GGATGGCTGT TAAGTGTGTA CACCGAACAC AGCAGGACAG ACCAGAGATT  1380
TGTTTTATGA AAGGTGCTTA CGAACAAGTA ATTAAGTACT GTACTACATA CCAGAGCAAA  1440
GGGCAGACCT TGACACTTAC TCAGCAGCAG AGAGATGTGT ACCAACAAGA GAAGGCACGC  1500
ATGGGCTCAG CGGGACTCAG AGTTCTTGCT TTGGCTTCTG GTCCTGAACT GGGACAGCTG  1560
ACATTTCTTG GCTTGGTGGG AATCATTGAT CCACCTAGAA CTGGTGTGAA AGAAGCTGTT  1620
ACAACACTCA TTGCCTCAGG AGTATCAATA AAAATGATTA CTGGAGATTC ACAGGAGACT  1680
GCAGTTGCAA TCGCCAGTCG TCTGGGATTG TATTCCAAAA CTTCCCAGTC AGTCTCAGGA  1740
GAAGAAATAG ATGCAATGGA TGTTCAGCAG CTTTCAGAAT AGTACCAAA GGTTGCAGTA  1800
TTTTACAGAG CTAGCCCAAG GCACAAGATG AAAATTATTA AGTCGCTACA GAAGAACGGT  1860
TCAGTTGTAG CCATGACAGG AGATGGAGTA AATGATGCAA TTGCTCTGAA GGCTGCAGAC  1920
ATTGGAGTTG CGATGGGCCA GACTGGTACA GATGTTTGCA AAGAGGCAGC AGACATGATC  1980
CTAGTGGATG ATGATTTTCA AACCATAATG TCTGCAATCG AAGAGGGTAA AGGGATTTAT  2040
```

```
AATAACATTA AAAATTTCGT TAGATTCCAG CTGAGCACGA GTATAGCAGC ATTAACTTTA  2100
ATCTCATTGG CTACATTAAT GAACTTTCCT AATCCTCTCA ATGCCATGCA GATTTTGTGG  2160
ATCAATATTA TTATGGATGG ACCCCCAGCT CAGAGCCTTG GAGTAGAACC AGTGGATAAA  2220
GATGTCATTC GTAAACCTCC TCGCAACTGG AAAGACAGCA TTTTGACTAA AAACTTGATA  2280
CTTAAAATAC TTGTTTCATC AATAATCATT GTTTGTGGGA CTTTGTTTGT CTTCTGGCGT  2340
GAGCTACGAG ACAATGTGAT TACACCTCGA GACACAACAA TGACCTTCAC ATGCTTTGTG  2400
TTTTTTGACA TGTTCAATGC ACTAAGTTCC AGATCCCAGA CCAAGTCTGT GTTTGAGATT  2460
GGACTCTGCA GTAATAGAAT GTTTTGCTAT GCAGTTCTTG GATCCATCAT GGGACAATTA  2520
CTAGTTATTT ACTTTCCTCC GCTTCAGAAG GTTTTTCAGA CTGAGAGCCT AAGCATACTG  2580
GATCTGTTGT TTCTTTTGGG TCTCACCTCA TCAGTGTGCA TAGTGGCAGA AATTATAAAG  2640
AAGGTTGAAA GGAGCAGGGA AAAGATCCAG AAGCATGTTA GTTCGACATC ATCATCTTTT  2700
CTTGAAGTAT GA
```

Gene name: ATPase, Ca++ transporting, type 2C, member 1
Unigene number: Hs.106778
Probeset Accession #:   N51919
Protein Accession #:   AAF27813
Signal sequence: none found
Transmembrane domains: 60-82, 86-108, 252-274, 282-304, 757-779, 827-849, 856-878
Pfam domains: Cation ATPase N [6-83], E1-E2 ATPase [89-324], Hydrolase [328-649], Cation ATPase C
[744-889]
Cellular Localization: not determined

```
1          11         21         31         41         51
|          |          |          |          |          |
MIPVLTSKKA SELPVSEVAS ILQADLQNGL NKCEVSHRRA FHGWNEFDIS EDEPLWKKYI    60
SQFKNPLIML LLASAVISVL MHQFDDAVSI TVAILIVVTV AFVQEYRSEK SLEELSKLVP   120
PECHCVREGK LEHTLARDLV PGDTVCLSVG DRVPADLRLF EAVDLSIDES SLTGETTPCS   180
KVTAPQPAAT NGDLASRSNI AFMGTLVRCG KAKGVVIGTG ENSEFGEVFK MMQAEEAPKT   240
PLQKSMDLLG KQLSPYSFGI IGIIMLVGWL LGKDILEMFT ISVSLAVAAI PEGLPIVVTV   300
TLALGVMRMV KKRAIVKKLP IVETLGCCNV ICSDKTGTLT KNEMTVTHIF TSDGLHAEVT   360
GVGYNQFGEV IVDGDVVHGF YNPAVSRIVE AGCVCNDAVI RNNTLMGKPT EGALIALAMK   420
MGLDGLQQDY IRKAEYPFSS EQKWMAVKCV HRTQQDRPEI CFMKGAYEQV IKYCTTYQSK   480
GQTLTLTQQQ RDVYQQEKAR MGSAGLRVLA LASGPELGQL TPLGLVGIID PPRTGVKEAV   540
TTLIASGVSI KMITGDSQET AVAIASRLGL YSKTSQSVSG EEIDAMDVQQ LSQIVPKVAV   600
FYRASPRHKM KIIKSLQKNG SVVAMTGDGV NDAVALKAAD IGVAMGQTGT DVCKEAADMI   660
LVDDDFQTIM SAIEEGKGIY NNIKNFVRFQ LSTSIAALTL ISLATIMNFP NPLNAMQILW   720
INIIMDGPPA QSLGVEPVDK DVIRKPPRNW KDSILTKNLI LKILVSSIII VCGTLFVFWR   780
ELRDNVITPR DTTMTFTCFV FFDMFNALSS RSQTESVPEI GLCSNRMPCY AVLGSIMGQL   840
LVIYFPPLQK VFQTESLSIL DLLFLLGLTS SVCIVAEIIK KVERSREKIQ KHVSSTSSSF   900
LEV
```

Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #:   N62096
Nucleic Acid Accession #:   N62096
Coding sequence: 1-1284 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CGGCCGCAGA GAGGATTGCC TTATTCAATG    60
AAGCAAGCTG GGTTTCCTTT GGGAATATTG CTTTTATTCT GGGTTTCATA TGTTACAGAC   120
TTTTCCCTTG TTTTATTGAT AAAAGGGAGG GCCCTCTCTG GAACAGATAC CTACCAGTCT   180
TTGGTCAATA AAACTTTCGG CTTTCCAGGG TATCTGCTCC TCTCTGTTCT TCAGTTTTTG   240
TATCCTTTTA TAGCAATGAT AAGTTACAAT ATAATAGCTG GAGATACTTT GAGCAAAGTT   300
TTTCAAAGAA TCCCAGGAGT TGATCCTGAA AACGTGTTTA TTGGTCGCCA CTTCATTATT   360
GGACTTTCCA CAGTTACCTT TACTCTGCCT TTATCCTTGT ACCGAAATAT AGCAAAGCTT   420
GGAAAGGTCT CCCTCATCTC TACAGGTTTA ACAACTCTGA TTCTTGGAAT TGTAATGGCA   480
AGGGCAATTT CACTGGGTCC ACACATACCA AAAACAGAAG ACGCTTGGGT ATTTGCAAAG   540
CCCAATGCCA TTCAAGCGGT CGGGGTTATG TCTTTTGCAT TTATTTGCCA CCATAACTCC   600
TTCTTAGTTT ACAGTTCTCT AGAAGAACCC ACAGTAGCTA AGTGGTCCCG CCTTATCCAT   660
ATGTCCATCG TGATTTCTGT ATTTATCTGT ATATTCTTTG CTACATGTGG ATACTTGACA   720
TTTACTGGCT TCACCCAAGG GGACTTATTT GAAAATTACT GCAGAAATGA TGACCTGGTA   780
ACATTTGGAA GATTTGTTA TGGTGTCACT GTCATTTTGA CATACCCTAT GGAATGGCTTT   840
GTGACAAGAG AGGTAATTGC CAATGTGTTT TTTGGTGGGA ATCTTTCATC GGTTTTCCAC   900
ATTGTTGTAA CAGTGATGGT CATCACTGTA GCCACGCTTG TGTCATTGCT GATTGATTGC   960
CTCGGGATAG TTCTAGAACT CAATGGTGTG CTCTGTGCAA CTCCCCTCAT TTTTATCATT  1020
CCATCAGCCT GTTATCTGAA ACTGTCTGAA GAACCAAGGA CACACTCCGA TAAGATGATG  1080
TCTTGTGTCA TGCTTCCCAT TGGTGCTGTG GTGATGGTTT TTGGATTCGT CATGGCTATT  1140
ACAAATACTC AAGACTGCAC CCATGGGCAG GAAATGTTCT ACTGCTTTCC TGACAATTTC  1200
TCTCTCACAA ATACCTCAGA GTCTCATGTT CAGCAGACAA CACAACTTTC TACTTTAAAT  1260
ATTAGTATCT TTCAACTCGA GTAA
```

Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #:   N62096
Protein Accession #:   none found
Signal sequence: none found
Transmembrane domains: 28-50, 66-88-112-134, 142-164, 217-239, 260-282, 298-320, 327-348, 359-381

Cellular Localization: plasma membrane

```
  1         11        21        31        41        51
  |         |         |         |         |         |
MGYQRQEPVI PPQRGLPYSM KQAGFPLGIL LLFWVSYVTD PSLVLLIKGG ALSGTDTYQS   60
LVNKTFGFPG YLLLSVLQFL YPFIAMISYN IIAGDTLSKV FQRIPGVDPE NVPIGRHPII  120
GLSTVTFTLP LSLYRNIAKL GKVSLISTGL TTLILGIVMA RAISLGPHIP KTEDAWVPAK  180
PNAIQAVGVM SFAFICHHNS PLVYSSLEEP TVAKWSRLIH MSIVISVPIC IFFATCGYLT  240
FTGFTQGDLF ENYCRNDDLV TFGRPCYGVT VILTYPMECF VTREVIANVF FGGNLSSVFH  300
IVVTVMVITV ATLVSLLIDC LGIVLELNGV LCATPLIFII PSACYLKLSE EPRTHSDKIM  360
SCVMLPIGAV VMVFGFVMAI TNTQDCTHGQ EMFYCFPDNF SLTNTSESHV QQTTQLSTLN  420
ISIFQLE
```

<u>A191 DNA SEQUENCE</u>
Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #:   N62096
Nucleic Acid Accession #: N62096
Coding sequence: 1-1203 (underlined sequences correspond to start and stop codons)

```
  1         11        21        31        41        51
  |         |         |         |         |         |
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGT TTTCCCTTGT TTTATTGATA   60
AAAGGAGGGG CCCTCTCTGG AACAGATACC TACCAGTCTT TGGTCAATAA AACTTTCGGC  120
TTTCCAGGGT ATCTGCTCCT CTCTGTTCTT CAGTTTTTGT ATCCTTTTAT AGCAATGATA  180
AGTTACAATA TAATAGCTGG AGATACTTTG AGCAAAGTTT TTCAAAGAAT CCCAGGAGTT  240
GATCCTGAAA ACGTGTTTAT TGGTCGCCAC TTCATTATTG GACTTTCCAC AGTTACCTTT  300
ACTCTGCCTT TATCCTTGTA CCGAAATATA GCAAAGGTCTG GAAAGGTCTC CCTCATCTCT  360
ACAGGTTTAA CAACTCTGAT TCTTGGAATT GTAATGGCAA GGGCAATTTC ACTGGGTCCA  420
CACATACCAA AAACAGAAGA CGCTTGGGTA TTTGCAAAGC CCAATGCCAT TCAAGCGGTC  480
GGGGTTATGT CTTTTGCATT TATTTGCCAC CATAACTCCT TCTTAGTTTA CAGTTCTCTA  540
GAAGAACCCA CAGTAGCTAA GTGGTCCCGC CTTATCCATA TGTCCATCGT GATTTCTGTA  600
TTTATCTGTA TATTCTTTGC TACATGTGGA TACTTGACAT TTACTGGCTT CACCCAAGGG  660
GACTTATTTG AAAATTACTG CAGAAATGAT GACCTGGTAA CATTTGGAAG ATTTTGTTAT  720
GGTGTCACTG TCATTTTGAC ATACCCTATG GAATGCTTTG TGACAAGAGA GGTAATTGCC  780
AATGTGTTTT TTGGTGGGAA TCTTTCATCG GTTTTCCACA TTGTTGTAAC AGTGATGGTC  840
ATCACTGTAG CCACGCTTGT GTCATTGCTG ATTGATTGCC TCGGGATAGT TCTAGAACTC  900
AATGGTGTGC TCTGTGCAAC TCCCCTCATT TTTATCATTC CATCAGCCTG TTATCTGAAA 960
CTGTCTGAAG AACCAAGGAC ACACTCCGAT AAGATTATGT CTTGTGTCAT GCTTCCCATT 1020
GGTGCTGTGG TGATGGTTTT TGGATTCGTC ATGGCTATTA CAAATACTCA AGACTGCACC 1080
CATGGGCAGG AAATGTTCTA CTGCTTTCCT GACAATTTCT CTCTCACAAA TACCTCAGAG 1140
TCTCATGTTC AGCAGACAAC ACAACTTTCT ACTTTAAATA TTAGTATCTT TCAACTCGAG 1200
TAA
```

<u>A192 Protein sequence:</u>
Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #:   N62096
Protein Accession #:   none found
Signal sequence: 1-26
Transmembrane domains: 45-61, 92-108, 117-132, 191-207, 274-290, 297-313, 335-351
Cellular Localization: plasma membrane

```
  1         11        21        31        41        51
  |         |         |         |         |         |
MGYQRQEPVI PPQFSLVLLI KGGALSGTDT YQSLVNKTFG PPGYLLLSVL QFLYPFIAMI   60
SYNIIAGDTL SKVPQRIPGV DPENVFIGRH FIIGLSTVTF TLPLSLYRNI AKLGKVSLIS  120
TGLTTLILGI VMARAISLGP HIPKTEDAWV FAKPNAIQAV GVMSFAFICH HNSFLVYSSL  180
EEPTVAKWSR LIHMSIVISV PICIFFATCG YLTFTGFTQG DLFENYCRND DLVTFGRFCY  240
GVTVILTYPM ECFVTREVIA NVPFGGNLSS VFHIVVTVMV ITVATLVSLL IDCLGIVLEL  300
NGVLCATPLI FIIPSACYLK LSEEPRTHSD KIMSCVMLPI GAVVMVFGFV MAITNTQDCT  360
HGQEMFYCFP DNPSLTNTSE SHVQQTTQLS TLNISIFQLE
```

<u>A193 DNA SEQUENCE</u>
Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #:   N62096
Nucleic Acid Accession #: N62096
Coding sequence: 1-1140 (underlined sequences correspond to start and stop codons)

```
  1         11        21        31        41        51
  |         |         |         |         |         |
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGG TCAATAAAAC TTTCGGCTTT   60
CCAGGGTATC TGCTCCTCTC TGTTCTTCAG TTTTTGTATC CTTTTATAGC AATGATAAGT  120
TACAATATAA TAGCTGGAGA TACTTTGAGC AAAGTTTTTC AAAGAATCCC AGGAGTTGAT  180
CCTGAAAACG TGTTTATTGG TCGCCACTTC ATTATTGGAC TTTCCACAGT TACCTTTACT  240
CTGCCTTTAT CCTTGTACCG AAATATAGCA AAGCTGGAA AGGTCTCCCT CATCTCTACA  300
GGTTTAACAA CTCTGATTCT TGGAATTGTA ATGGCAAGGG CAATTTCACT GGGTCCACAC  360
ATACCAAAAA CAGAAGACGC TTGGGTATTT GCAAAGCCCA ATGCCATTCA AGCGGTCGGG  420
```

```
GTTATGTCTT TTGCATTTAT TTGCCACCAT AACTCCTTCT TAGTTTACAG TTCTCTAGAA    480
GAACCCACAG TAGCTAAGTG GTCCCGCCTT ATCCATATGT CCATCGTGAT TTCTGTATTT    540
ATCTGTATAT TCTTTGCTAC ATGTGGATAC TTGACATTTA CTGGCTTCAC CCAAGGGGAC    600
TTATTTGAAA ATTACTGCAG AAATGATGAC CTGGTAACAT TTGGAAGATT TTGTTATGGT    660
GTCACTGTCA TTTTGACATA CCCTATGGAA TGCTTTGTGA CAAGAGAGGT AATTGCCAAT    720
GTGTTTTTTG GTGGGAATCT TTCATCGGTT TTCCACATTG TTGTAACAGT GATGGTCATC    780
ACTGTAGCCA CGCTTGTGTC ATTGCTGATT GATTGCCTCG GGATAGTTCT AGAACTCAAT    840
GGTGTGCTCT GTGCAACTCC CCTCATTTTT ATCATTCCAT CAGCCTGTTA TCTGAAACTG    900
TCTGAAGAAC CAAGGACACA CTCCGATAAG ATTATGTCTT GTGTCATGCT TCCCATTGGT    960
GCTGTGGTGA TGGTTTTTGG ATTCGTCATG GCTATTACAA ATACTCAAGA CTGCACCCAT   1020
GGGCAGGAAA TGTTCTACTG CTTTCCTGAC AATTTCTCTC TCACAAATAC CTCAGAGTCT   1080
CATGTTCAGC AGACAACACA ACTTTCTACT TTAAATATTA GTATCTTTCA ACTCGAGTAA
```

A194 Protein sequence:
Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #:   N62096
Protein Accession #:   none found
Signal sequence: none found
Transmembrane domains: 24-40, 70-86, 95-111, 171-186, 253-269, 276-292, 314-330
Cellular Localization: plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
|          |          |          |          |          |
MGYQRQEPVI PPQVNKTFGP PGYLLLSVLQ FLYPFIAMIS YNIIAGDTLS KVFQRIPGVD    60
PENVPIGRHF IIGLSTVTFT LPLSLYRNIA KLGKVSLIST GLTTLILGIV MARAISLGPH   120
IPKTEDAVVP AKPNAIQAVG VMSFAFICHH NSFLVYSSLE EPTVAKWSRL IHMSIVISVF   180
ICIFFATCGY LTPTGPTQGD LFENYCRNDD LVTPGRFCYG VTVILTYPME CFVTREVIAN   240
VFFGGNLSSV FHIVVTVMVI TVATLVSLLI DCLGIVLELN GVLCATPLIF IIPSACYLKL   300
SEEPRTHSDK IMSCVMLPIG AVVMVFGFVM AITNTQDCTH GQEMFYCFPD NFSLTNTSES   360
HVQQTTQLST LNISIFQLE
```

A195 DNA SEQUENCE:
Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #:   N62096
Nucleic Acid Accession #: N62096
Coding sequence: 1-1389 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
|          |          |          |          |          |
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGA GAGATTTAGA TGACAGAGAA    60
ACCCTTGTTT CTGAACATGA GTATAAAGAG AAAACCTGTC AGTCTGCTGC TCTTTTTAAT   120
GTTGTCAACT CGATTATAGG ATCTGGTATA ATAGGATTGC CTTATTCAAT GAAGCAAGCT   180
GGGTTTCCTT TGGGAATATT GCTTTTATTC TGGGTTTCAT ATGTTACAGA CTTTTCCCTT   240
GTTTTATTGA TAAAAGGAGG GGCCCTCTCT GGAACAGATA CCTACCAGTC TTTGGTCAAT   300
AAAACTTTCG GCTTTCCAGG GTATCTGCTC CTCTCTGTTC TTCAGTTTTT GTATCCTTTT   360
ATAGCAATGA TAAGTTACAA TATAATAGCT GGAGATACTT TGAGCAAAGT TTTTCAAAGA   420
ATCCCAGGAG TTGATCCTGA AAACGTGTTT ATTGGTCGCC ACTTCATTAT TGGACTTTCC   480
ACAGTTACCT TTACTCTGCC TTTATCCTTG TACCCGAAATA TAGCAAAGCT TGGAAAGGTC   540
TCCCTCATCT CTACAGGTTT AACAACTCTG ATTCTTGGAA TTGTAATGGC AAGGGCAATT   600
TCACTGGGTC CACACATACC AAAAACAGAA GACGCTTGGG TATTTGCAAA GCCCAATGCC   660
ATTCAAGCGG TCGGGGTTAT GTCTTTTGCA TTTATTTGCC ACCATAACTC CTTCTTAGTT   720
TACAGTTCTC TAGAAGAACC ACACAGTAGCT AAGTGGTCCC GCCTTATCCA TATGTCCATC   780
GTGATTTCTG TATTTATCTG TATATTCTTT GCTACATGTG GATACTTGAC ATTTACTGGC   840
TTCACCCAAG GGGACTTATT TGAAAATTAC TGCAGAAATG ATGACCTGGT AACATTTGGA   900
AGATTTGTT ATGGTGTCAC TGTCATTTTG ACATACCCTA TGGAATGCTT TGTGACAAGA   960
GAGGTAATTG CCAATGTGTT TTTTGGTGGG AATCTTTCCA CATTGTTGTC                 1020
ACAGTGATGG TCATCACTGT AGCCACGCTT GTGTCATTGC TGATTGATTG CCTCGGGATA  1080
GTTCTAGAAC TCAATGGTGT GCTCTGTGCA ACTCCCCTCA TTTTTATCAT TCCATCAGCC  1140
TGTTATCTGA AACTGTCTGA AGAACCAAGG ACACACTCCG ATAAGATTAT GTCTTGTGTC  1200
ATGCTTCCCA TTGGTGCTGT GGTGATGGTT TTTGGATTCG TCATGGCTAT TACAAATACT  1260
CAAGACTGCA CCCATGGGCA GGAAATGTTC TACTGCTTTC CTGACAATTT CTCTCTCACA  1320
AATACCTCAG AGTCTCATGT TCAGCAGACA ACACAACTTT CTACTTTAAA TATTAGTATC  1380
TTTCAATGA
```

A196 Protein sequence:
Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #:   N62096
Protein Accession #:   none found
Signal sequence: none found
Transmembrane domains: 37-53, 66-82, 109-125, 155-172, 180-196, 255-271, 338-354, 361-377, 399-415
Cellular Localization: plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
MGYQRQEPVI PPQRDLDDRE TLVSEHEYKE KTCQSAALFN VVNSIIGSGI IGLPYSMKQA    60
```

```
     GFPLGILLLF WVSYVTDPSL VLLIKGGALS GTDTYQSLVN KTFGFPGYLL LSVLQFLYPF  120
     IAMISYNIIA GDTLSKVPQR IPGVDPENVF IGRHPIIGLS TVTFTLPLSL YRNIAKLGKV  180
  5  SLISTGLTTL ILGIVMARAI SLGPHIPKTE DAWVFAKPNA IQAVGVMSPA FICHHNSFLV  240
     YSSLEEPTVA KWSRLIHMSI VISVPICIFF ATCGYLTFTG FTQGDLFENY CRNDDLVTFG  300
     RFCYGVTVIL TYPMECFVTR EVIANVFFGG NLSSVPHIVV TVMVITVATL VSLLIDCLGI  360
     VLELNGVLCA TPLIFIIPSA CYLKLSEEPR THSDKIMSCV MLPIGAVVMV FGFVMAITNT  420
     QDCTHGQEMP YCPPDMFSLT NTSESHVQQT TQLSTLNISI FQ
```

```
     1         11        21        31        41        51
     |         |         |         |         |         |
     ATGGAGGATG CCTTCGGGGC AGCCGTGGTG ACCGTGTGGG ACAGCGATGC ACACACCACG   60
     GAGAAGCCCA CCGATGCCTA CGGAGAGCTG GACTTCACCG GGGCCGGCCG CAAGCACAGC  120
     AATTTCCTCC GGCTCTCTGA CCGAACGGAT CCAGCTGCAG TTTATAGTCT GGTCACACGC  180
     ACATGGGGCT TCCGTGCCCC GAACCTGGTG GTGTCAGTGC TGGGGGGATC GGGGGGCCCC  240
     GTCCTCCAGA CCTGGCTGCA GGACCTGCTG CGTCGTGGGC TGGTGCGGGC TGCCCAGAGC  300
     ACAGGAGCCT GGATTGTCAC TGGGGGTCTG CACACGGGCA TCGGCCGGCA TGTTGGTGTG  360
     GCTGTACGGG ACCATCAGAT GGCCAGCACT GGGGGCACCA AGGTGGTGGC CATGGGTGTG  420
     GCCCCCTGGG GTGTGGTCCG GAATAGAGAC ACCCTCATCA ACCCCAAGGG CTCGTTCCCT  480
     GCGAGGTACC GGTGGCGCGG TGACCCGGAG GACGGGGTCC AGTTTCCCCT GGACTACAAC  540
     TACTCGGCCT TCTTCCTGGT GGACGACGGC ACACACGGCT GCCTGGGGGG CGAGAACCGC  600
     TTCCGCTTGC GCCTGGAGTC CTACATCTCA CAGCAGAAGA CGGGCGTGGG AGGGACTGGA  660
     ATTGACATCC CTGTCCTGCT CCTCCTGATT GATGGTGATG AGAAGATGTT GACGCGAATA  720
     GAGAACGCCA CCCAGGCTCA GCTCCCCATGT CTCCTCGTGG CTGGCTCAGG GGGAGCTGCG  780
     GACTGCCTGG CGGAGACCCT GGAAGACACT CTGGCCCCAG GGAGTGGGGG AGCCAGGCAA  840
     GGCGAAGCCC GAGATCGAAT CAGGCGTTTC TTTCCCAAAG GGGACCTTGA GGTCCTGCAG  900
     GCCCAGGTGG AGAGGATTAT GACCCGGAAG GAGCTCCTGA CAGTCTATTC TTCTGAGGAT  960
     GGGTCTGAGG AATTCGAGAC CATAGTTTTG AAGGCCCTTG TGAAGGCCTG TGGGAGCTCG 1020
     GAGGCCTCAG CCTACCTGGA TGAGCTGCGT TTGGCTGTGG CTTGGAACCG CGTGGACATT 1080
     GCCCAGAGTG AACTCTTTCG GGGGGACATC CAATGGCGGT CCTTCCATCT CGAAGCTTCC 1140
     CTCATGGACG CCCTGCTGAA TGACCGGCCT GAGTTCGTGC GCTTGCTCAT TTCCCACGGC 1200
     CTCAGCCTGG GCCACTTCCT GACCCCGATG CGCCTGGCCC AACTCTACAG GGGGGACATC 1260
     TCCAACTCGC TCATCCGCAA CCTTTTGGAC CAGGCGTCCC ACAGCGCAGG CACCAAAGCC 1320
     CCAGCCCTAA AAGGGGGAGC TGCGGAGCTC CGGCCCCCTG ACGTGGGGCA TGTGCTGAGG 1380
     ATGCTGCTGG GGAAGATGTG CGCGCCGAGG TACCCCTCCG GGGGCGCCTG GGACCCTCAC 1440
     CCAGGCCAGG GCTTCGGGGA GAGCATGTAT CTGCTCTCGG ACAAGGCCAC CTCGCCGCTC 1500
     TCGCTGGATG CTGGCCTCGG GCAGGCCCCC TGGAGCGACC TGCTTCTTTG GGCACTGTTG 1560
     CTGAACAGGG CACAGATGGC CATGTACTTC TGGGAGATGG GTTCCAATGC AGTTTCCTCA 1620
     GCTCTTGGGG CCTGTTTGCT GCTCCGGGTG ATGGCACGGC TGGAGCCTGA CGCTGAGGAG 1680
     GCAGCACGGA GGAAAGACCT GGCCGTTCAAG TTGAGGGGA TGGGCGTTGA CCTCTTTGGC 1740
     GAGTGCTATC GCAGCAGTGA GGTGAGGGCT GCCCGCCTCC TCCTCCGTCG CTGCCCGCTC 1800
     TGGGGGGATG CCACTTGCCT CCAGCTGGCC ATGCAAGCTG ACGCCGGTGC CTTCTTTGCC 1860
     CAGGATGGGG TACAGTCTCT GCTGACACAG AAGTGGTGGG GAGATATGGC CAGCACTACA 1920
     CCCATCTGGG CCCTGGTTCT CGGCCTTCTT TGCCCTCCAC TCATCTACAC CCGCCTCATC 1980
     ACCTTCAGGA AATCAGAAGA GGAGCCCACA CGGGAGGAGC TAGAGTTTGA CATGGATAGT 2040
     GTCATTAATG GGGAAGGGCC TGTCGGGACG GCGGACCCAG CCGAGAAGAC GCCGGCTGGG 2100
     GTCCCGCGCC AGTCGGGCCG TCCGGGTTGC TGCGGGGGCC GCTGCGGGGG GCGGCCGGTGC 2160
     CTACGCCGCT GGTTCCACTT CTGGGGCGCG CCGGTGACCA TCTTCATGGG CAACGTGGTC 2220
     AGCTACCTGC TGTTCCTGCT GCTTTTCTCG CGGGTGCTGC TCGTGGATTT CCAGCCGGCG 2280
     CCGCCCGGCT CCCTGGAGCT GCTGCTCTAT TTCTGGGCTT TCACGCTGCT GTGCGAGGAA 2340
     CTGCGCCAGG GCCTGAGCGG AGGCGGGGGC AGCCTGGCCA GCGGGGGCCC CGGGCCTGGC 2400
     CATGCCTCAC TGAGCCAGCG CCTGCGCCTC TACCTGCCCG ACAGCTGGAA CCAGTGCGAC 2460
     CTAGTGGCTC TCACCTGCTT CCTCCTGGGC GTGGGCTGCC GGCTGACCCC GGGTTTGTAC 2520
     CACCTGGGCC GCACTGTCCT CTGCCATCGAC TTCATGGTTT TCACGGTGCG GCTCGTTCAC 2580
     ATCTTCACGG TCAACAAACA GCTGGGGCCC AAGATCGTCA TCGTGAGCAA GATGATGAAG 2640
     GACGTGTTCT TCTTCCTCTT CTTCCTCGGC GTGTGGCTGG TAGCCTATGG CGTGGCCACG 2700
     GAGGGGCTCC TGAGGCCACG GGACAGTGAC TTCCCAGTA TCCTGCGCCG CGTCTTCTAC 2760
     CGTCCCTACC TGCAGATCTT CGGGGCAGATT CCCCAGGAGG ACATGGACGT GGCCCTCATG 2820
     GAGCACAGCA ACTGCTCGTC GGAGCCCGGC TTCTGGGCAC ACCCTCCTGG GGCCCAGGCG 2880
     GGCACCTGCG TCTCCCAGTA TGCCAACTGG CTGGTGGTGC TGCTCCTCGT CATCTTCCTG 2940
     CTCGTGGCCA ACATCCTGCT GGTCAACTTG CTCATTGCCA TGTTCAGTTA CACATTCGGC 3000
     AAAGTACAGG GCAACAGCGA TCTCTACTGG AAGGCGCAGC GTTACCGGCCT CATCCGGGAA 3060
     TTCCACTCTC GGCCCGCGCT GGCCCCGCCC TTTATCGTCA TCTCCCACTT GCGCCTCCTG 3120
     CTCAGGCAAT TGTGCAGGCG ACCCCGGGAG CCCCAGCCGT CCTCGCCGGC CCTCGAGCAT 3180
     TTCCGGGTTT ACCTTTCTAA GGAAGCCGAG CGGAAGCTGC TAACGTGGGA ATCGGTGCAT 3240
     AAGGAGAACT TTCTGCTGGC ACGCGCTAGG GACAAGCGGG AGAGCGACTC CGAGCGTCTG 3300
     AAGCGCACGT CCCAGAAGGT GGACTTGGCA CTGAAACAGC TGGGACACAT CCGCGGAGTAC 3360
     GAACAGCGCC TGAAAGTGCT GGAGCGGGAG GTCCAGCAGT GTAGCCGCGT CCTGGGGTGG 3420
     GTGGCCGAGG CCCTGAGCCG CTCTGCCTTG CTGCCCCCAG GTGGGCCGCC ACCCCCTGAC 3480
     CTGCCTGGGT CCAAAGACTG A                                           3480
```

Signal sequence: none found
Transmembrane domains: 214-230, 537-556, 642-662, 730-752, 760-782, 815-837, 842-864, 877-899, 973-995
Cellular Localization: plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
MEDAFGAAVV TVWDSDAHTT EKPTDAYGEL DFTGAGRKHS NFLRLSDRTD PAAVYSLVTR    60
TWGFRAPNLV VSVLGGSGGP VLQTWLQDLL RRGLVRAAQS TGAWIVTGGL HTGIGRHVGV   120
AVRDHQMAST GGTKVVAMGV APWGVVRNRD TLINPKGSFP ARYRWRGDPE DGVQPPLDYN   180
YSAFFLVDDG THGCLGGENR FRLRLESYIS QQKTGVGGTG IDIPVLLLLI DGDEKMLTRI   240
ENATQAQLPC LLVAGSGGAA DCLAETLEDT LAPGSGGARQ GEARDRIRRF FPKGDLEVLQ   300
AQVERIMTRK ELLTVYSSED GSEEPETIVL KALVKACGSS EASAYLDELR LAVAWNRVDI   360
AQSELFRGDI QWRSFHLEAS LMDALLNDRP EFVRLLISHG LSLGHFLTFM RLAQLYSAAP   420
SNSLIRNLLD QASHSAGTKA PALKGGAAEL RPPDVGHVLR MLLGKMCAPR YPSGGAWDPH   480
PGQGFGESMY LLSDKATSPL SLDAGLGQAP WSDLLLWALL LNRAQMAMYP WEMGSNAVSS   540
ALGACLLLRV MARLEPDAEE AARRKDLAFK FEGMGVDLFG ECYRSSEVRA ARLLLRRCPL   600
WGDATCLQLA MQADARAFFA QDGVQSLLTQ KWHGDMASTT PIWALVLAFF CPPLIYTRLI   660
TFRKSEEEPT REELEFDMDS VINGEGPVGT ADPAEKTPLG VPRQSGRPGC CGGRCGGRRC   720
LRRWFHFWGA PVTIFMGNVV SYLLFLLLFS RVLLVDFQPA PPGSLELLLY FWAFTLLCEE   780
LRQGLSGGGG SLASGGPGPG HASLSQRLRL YLADSWNQCD LVALTCFLLG VGCRLTPGLY   840
HLGRTVLCID FMVFTVRLLH IFTVNKQLGP KIVIVSKMMK DVFFFLFFLG VWLVAYGVAT   900
EGLLRPRDSD FPSILRRVFY RPYLQIFGQI PQEDMDVALM ERSNCSSEPQ FWAHPPGAQA   960
GTCVSQYANW LVVLLLVIFL LVANILLVNL LIAMFSYTFG KVQGNSDLYW KAQRYRLIRE  1020
PHSRPALAPP FIVISHLRLL LRQLCRRPRS PQPSSPALEH FRVYLSKEAE RKLLTWESVH  1080
KENFLLARAR DKRESDSERL RRTSQKVDLA LKQLGHIRBY EQRLKVLERE VQQCSRVLGW  1140
VAEALSRSAL LPPGGPPPPD LPGSKD
```

A199 DNA SEQUENCE

Gene name: ESTs
Unigene number: Hs.40808
Probeset Accession #:   AA054237
Nucleic Acid Accession #:   AA054237
Coding sequence: 1-894 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGAGCCGC GGGCGCTCGT CACGGCGCTC AGCCTCGGCC TCAGCCTGTG CTCCCTGGGG    60
CTGCTCGTCA CGGCCATCTT CACCGACCAC TGGTACGAGA CCGACCCCCG GCGCCACAAG   120
GAGAGCTGCG AGCGCAGCCG CGCGGGCGCC GACCCCCCGG ACCAGAAGAA CCGCCTGATG   180
CCGCTGTCGC ACCTGCCGCT GCGGGACTCG CCCCCGCTGG GGCGCCGGCT GCTCCCGGGC   240
GGCCCGGGGC GCGCCGACCC CGAGTCCTGG CGCTCGCTCC TGGGGCTCGG CGGGCTGGAC   300
GCCGAGTGCG GCCGGCCCCT CTTCGCCACC TACTCGGGCC TCTGGAGGAA GTGCTACTTC   360
CTGGGCATCG ACGGGGACAT CGACACCCTC ATCCTGAAAG GTATTGCGCA GCGATGCACG   420
GCCATCAAGT ACCACTTTTC TCAGCCCATC CGCTTGACGA ACATTCCTTT TAATTTAACC   480
AAGACCATAC AGCAAGATGA GTGGCACCTG CTTCATTTAA GAAGAATCAC TGCTGGCTTC   540
CTCGGCATGG CCGTAGCCGT CCTTCTCTGC GGCTGCATTG TGGCCACAGT CAGTTTCTTC   600
TGGGAGGAGA GCTTGACCCA GCACGTGGCT GGACTCCTGT TCCTCATGAC AGGGATATTT   660
TGCACCATTT CCCTCTGTAC TTATGCCGCC AGTATCTCGT ATGATTTGAA CCGGCTCCCA   720
AAGCTAATTT ATAGCCTGCC TGCTGATGTG GAACATGGTT ACAGCTGGTC CATCTTTTGC   780
GCCTGGTGCA GTTTAGGCTT TATTGTGGCA GCTGGAGGTC TCTGCATCGC TTATCCGTTT   840
ATTAGCCGGA CCAAGATTGC ACAGCTAAAG TCTGGCAGAG ACTCCACGGT ATGA
```

A200 Protein sequence:

Gene name: ESTs
Unigene number: Hs.40808
Probeset Accession #:   AA054237
Protein Accession #:   none found
Signal sequence: 1-18
Transmembrane domains: 179-201, 209-231, 257-279
Cellular Localization: plasma membrane

```
1          11         21         31         41         51
|          |          |          |          |          |
MEPRALVTAL SLGLSLCSLG LLVTAIFTDH WYETDPRRHK ESCERSRAGA DPPDQKNRLM    60
PLSHLPLRDS PPLGRRLLPG GPGRADPESW RSLLGLGGLD AECGRPLFAT YSGLWRKCYF   120
LGIDRDIDTL ILKGIAQRCT AIKYHFSQPI RLRNIPFNLT KTIQQDEWHL LHLRRITAGF   180
LGMAVAVLLC GCIVATVSFF WEESLTQHVA GLLFLWTGIF CTISLCTYAA SISYDLNRLP   240
KLIYSLPADV EHGYSWSIFC AWCSLGFIVA AGGLCIAYPF ISRTKIAQLK SGRDSTV
```

A201 DNA SEQUENCE

Gene name: CDA14
Unigene number: Hs.26813
Probeset Accession #:   N32912
Nucleic Acid Accession #:   NM_016570
Coding sequence: 1-1134 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAGGCGAC TGAATCGGAA AAAAACTTTA AGTTTGGTAA AAGAGTTGGA TGCCTTTCCG    60
AAGGTTCCTG AGAGCTATGT AGAGACTTCA GCCAGTGGAG GTACAGTTTC TCTAATAGCA   120
```

```
        TTTACAACTA TGGCTTTATT AACCATAATG GAATTCTCAG TATATCAAGA TACATGGATG    180
        AAGTATGAAT ACGAAGTAGA CAAGGATTTT TCTAGCAAAT TAAGAATTAA TATAGATATT    240
        ACTGTTGCCA TGAAGTGTCA ATATGTTGGA GCGGATGTAT TGGATTTAGC AGAAACAATG    300
        GTTGCATCTG CAGATGGTTT AGTTTATGAA CCAACAGTAT TTGATCTTTC ACCACAGCAG    360
        AAAGAGTGGC AGAGGATGCT GCAGCTGATT CAGAGTAGGC TACAAGAAGA GCATTCACTT    420
        CAAGATGTGA TATTTAAAAG TGCTTTTAAA AGTACATCAA CAGCTCTTCC ACCAAGAGAA    480
        GATGATTCAT CACAGTCTCC AAATGCATGC AGAATTCATG GCCATCTATA TGTCAATAAA    540
        GTAGCAGGGA ATTTTCACAT AACAGTGGGC AAGGCAATTC CACATCCTCG TGGTCATGCA    600
        CATTTGGCAG CACTTGTCAA CCATGAATCT TACAATTTTT CTCATAGAAT AGATCATTTG    660
        TCTTTTGGAG AGCTTGTTCC AGCAATTATT AATCCTTTGA ATGGAACTGA AAAAATTGCT    720
        ATAGATCACA ACCAGATGTT CCAATATTTT ATTACAGTTG TGCCAACAAA ACTACATACA    780
        TATAAAATAT CAGCAGACAC CCATCAGTTT TCTGTGACAG AAAGGGAACG TATCATTAAC    840
        CATGCTGCAG GCAGCCATGG AGTCTCTGGG ATATTTATGA AATATGATCT CAGTTCTCTT    900
        ATGGTGACAG TTACTGAGGA GCACATGCCA TTCTGGCAGT TTTTTGTAAG ACTCTGTGGT    960
        ATTGTTGGAG GAATCTTTTC AACAACAGGC ATGTTACATG GAATTGGAAA ATTTATAGTT   1020
        GAAATAATTT GCTGTCGTTT CAGACTTGGA TCCTATAAAC CTGTCAATTC TGTTCCTTTT   1080
        GAGGATGGCC ACACAGACAA CCACTTACCT CTTTTAGAAA ATAATACACA TTGA
```

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        MRRLNRKKTL SLVKELDAFP KVPESYVETS ASGGTVSLIA FTTMALLTIM EFSVYQDTWM    60
        KYEYEVDKDF SSKLRINIDI TVAMKCQYVG ADVLDLAETM VASADGLVYE PTVFDLSPQQ   120
        KEWQRMLQLI QSRLQEEHSL QDVIFKSAFK STSTALPPRE DDSSQSPNAC RIHGHLYVNK   180
        VAGNFHITVG KAIPHPRGHA HLAALVNHES YNFSHRIDHL SFGELVPAII NPLDGTEKIA   240
        IDHNQMFQYF ITVVPTKLHT YKISADTHQF SVTERRIIN HAAGSHGVSG IFMKYDLSSL   300
        MVTVTEEHMP FWQFFVRLCG IVGGIFSTTG MLHGIGKFIV EIICCRFRLG SYKPVNSVPF   360
        EDGHTDNHLP LLENNTH
```

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        CGCGCGTCTC CGCCGCCCGC GTGACTTCTG CCTGCGCTCC TTCTCTGAAC GCTCACTTCC    60
        GAGGAGACGC CGACGATGAA GACACCGTGG AAGATTCTTC TGGGACTGCT GGGTGCTGCT   120
        GCGCTTGTCA CCATCATCAC CGTGCCCGTG GTTCTGCTGA ACAAAGGCAC AGATGATGCT   180
        ACAGCTGACA GTCGCAAAAC TTACACTCTA ACTGATTACT TAAAAAATAC TTATAGACTG   240
        AAGTTATACT CCTTAAGATG GATTTCAGAT CATGAATATC TCTACAAACA AGAAAATAAT   300
        ATCTTGGTAT TCAATGCTGA ATATGGAAAC AGCTCAGTTT TCTTGGAGAA CAGTACATTT   360
        GATGAGTTTG GACATTCTAT CAATGATTAT TCAATATCTC CTGATGGGCA GTTTATTCTC   420
        TTAGAATACA ACTACGTGAA GCAATGGAGG CATTCCTACA CAGCTTCATA TGACATTTAT   480
        GATTTAAATA AAAGGCAGCT GATTACAGAA GAGAGGATTC CAAACAACAC ACAGTGGGTC   540
        ACATGGTCAC CAGTGGGTCA TAAATTGGCA TATGTTTGGA ACAATGACAT TTATGTTAAA   600
        ATTGAACCAA ATTTACCAAG TTACAGAATC ACATGGACGG GGAAAGAAGA TATAATATAT   660
        AATGGAATAA CTGACTGGGT TTATGAAGAG GAAGTCTTCA GTGCCTACTC TGCTCTGTGG   720
        TGGTCTCCAA ACGGCACTTT TTTAGCATAT GCCCAATTTA ACGACACAGA AGTCCCACTT   780
        ATTGAATACT CCTTCTACTC TGATGAGTCA CTGCAGTACC CAAAGACTGT ACGGGTTCCA   840
        TATCCAAAGG CAGGAGCTGT GAATCCAACT GTAAAGTTCT TTGTTGTAAA TACAGACTCT   900
        CTCAGCTCAG TCACCAATGC AACTTCCATA CAAATCACTG CTCCTGCTTC TATGTTGATA   960
        GGGGATCACT ACTTGTGTGA TGTGACATGG GCAACACAAG AAAGAATTTC TTTGCAGTGG   1020
        CTCAGGAGGA TTCAGAACTA TTCGGTCATG GATATTTGTG ACTATGATGA ATCCAGTGGA   1080
        AGATGGAACT GCTTAGTGGC ACGGCAACAC ATTGAAATGA GTACTACTGG CTGGGTTGGA   1140
        AGATTTAGGC CTTCAGAACC TCATTTTACC CTTGATGGTA ATAGCTTCTA CAAGATCATC   1200
        AGCAATGAAG AAGGTTACAG ACACATTTGC TATTTCCAAA TAGATAAAAA AGACTGCACA   1260
        TTTATTACAA AAGGCACCTG GGAAGTCATC GGGATAGAAG CTCTAACCAG TGATTATCTA   1320
        TACTACATTA GTAATGAATA TAAAGGAATG CCAGGAGGAA GGAATCTTTA TAAAATCCAA   1380
        CTTATTGACT ATACAAAAGT GACATGCCTC AGTTGTGAGC TGAATCCGGA AAGGTGTCAG   1440
        TACTATTCTG TGTCATTCAG TAAAGAGGCG AAGTATTATC AGCTGAGATG TTCCGGTCCT   1500
        GGTCTGCCCC TCTATACTCT ACACAGCAGC GTGAATGATA AAGGGCTGAG AGTCCTGGAA   1560
        GACAATTCAG CTTTGGATAA AATGCTGCAG AATGTCCAGA TGCCCTCCAA AAAACTGGAC   1620
        TTCATTATTT TGAATGAAAC AAAATTTTGG TATCAGATGA TCTTGCCTCC TCATTTTGAT   1680
        AAATCCAAGA AATATCCTCT ACTATTAGAT GTGTATGCAG GCCCATGTAG TCAAAAAGCA   1740
        GACACTGTCT TCAGACTGAA CTGGGCCACT TACCTTGCAA CAGCAGAAAA CATTATAGTA   1800
        GCTAGCTTTG ATGGCAGAGG AAGTGGTTAC AAGGAGATA AGATCATGCA TGCAATCAAC   1860
        AGAAGACTGG GAACATTTGA AGTTGAAGAT CAAATTGAAG CAGCCAGACA ATTTTCAAAA   1920
        ATGGGATTTG TGGACAACAA ACGAATTGCA ATTTGGGGCT GGTCATATGG AGGGTACGTA   1980
        ACCTCAATGG TCCTGGGATC GGGAAGTGGC GTGTTCAAGT GTGGAATAGC CGTGGCGCCT   2040
```

```
GTATCCCGGT GGGAGTACTA TGACTCAGTG TACACAGAAC GTTACATGGG TCTCCCAACT 2100
CCAGAAGACA ACCTTGACCA TTACAGAAAT TCAACAGTCA TGAGCAGAGC TGAAAATTTT 2160
AAACAAGTTG AGTACCTCCT TATTCATGGA ACAGCAGATG ATAACGTTCA CTTTCAGCAG 2220
TCAGCTCAGA TCTCCAAAGC CCTGGTCGAT GTTGGAGTGG ATTTCCAGGC AATGTGGTAT 2280
ACTGATGAAG ACCATGGAAT AGCTAGCAGC ACAGCACACC AACATATATA TACCCACATG 2340
AGCCACTTCA TAAAACAATG TTTCTCTTTA CCTTAGCACC TCAAAATACC ATGCCATTTA 2400
AAGCTTATTA AAACTCATTT TTGTTTTCAT TATCTCAAAA CTGCACTGTC AAGATGATGA 2460
TGATCTTTAA AATACACACT CAAATCAAGA AACTTAAGGT TACCTTTGTT CCCAAATTTC 2520
ATACCTATCA TCTTAAGTAG GGACTTCTGT CTTCACAACA GATTATTACC TTACAGAAGT 2580
TTGAATTATC CGGTCGGGTT TTATTGTTTA AAATCATTTC TGCATCAGCT GCTGAAACAA 2640
CAAATAGGAA TTGTTTTTAT GGAGGCTTTG CATAGATTCC CTGAGCAGGA TTTTAATCTT 2700
TTTCTAACTG GACTGGTTCA AATGTTGTTC TCTTCTTTAA AGGGATGGCA AGATGTGGGC 2760
AGTGATGTCA CTAGGGCAGG GACAGGATAA GAGGGATTAG GGAGAGAAGA TAGCAGGGCA 2820
TGGCTGGGAA CCCAAGTCCA AGCATACCAA CACGAGCAGG CTACTGTCAG CTCCCCTCGG 2880
AGAAGAGCTG TTCACCACGA GACTGGCACA GTTTTCTGAG AAAGACTATT CAAACAGTCT 2940
CAGGAAATCA AATATCGAAA GCACTGACTT CTAAGTAAAC CACAGCAGTT GAAAGACTCC 3000
AAAGAAATGT AAGGGAAACT GCCAGCAACG CAGCCCCCAG GTGCCAGTTA TGGCTATAGG 3060
TGCTACAAAA ACACAGCAAG GGTGATGGGA AAGCATTGTA AATGTGCTTT TAAAAAAAAA 3120
TACTGATGTT CCTAGTGAAA GAGGCAGCTT GAAACTGAGA TGTGAACACA TCAGCTTGCC 3180
CTGTTAAAAG ATGAAAATAT TTGTATCACA AATCTTAACT TGAAGGAGTC CTTGCATCAA 3240
TTTTTCTTAT TTCATTTCTT TGAGTGTCTT AATTAAAAGA ATATTTTAAC TTCCTGGAC 3300
TCATTTTAAA AAATGGAACA TAAAATACAA TGTTATGTAT TATTATTCCC ATTCTACATA 3360
CTATGGAATT TCTCCCAGTC ATTTAATAAA TGTGCCTTCA TTTTTTC
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MKTPWKILLG LLGAAALVTI ITVPVVLLNK GTDDATADSR KTYTLTDYLK NTYRLKLYSL   60
RWISDHEYLY KQENNILVFN AEYGNSSVPL ENSTFDEFGH SINDYSISPD GQFILLSYNY  120
VKQWRHSYTA SYDIYDLNKR QLITEERIPN NTQWVTWSPV GHKLAYVWNN DIYVKIEPNL  180
PSYRITWTGK EDIIYNGITD WVYEEEVFSA YSALWWSPNG TFLAYAQFND TEVPLIEYSF  240
YSDESLQYPK TVRVPYPKAG AVNPTVKFFV VNTDSLSSVT NATSIQITAP ASMLIGDHYL  300
CDVTWATQER ISLQWLRRIQ NYSVMDICDY DESSGRWNCL VARQHIEMST TGWVGRFRPS  360
EPHFTLDGNS FYKIISNEEG YRHICYFQID KKDCTFITKG TWEVIGIEAL TSDYLYYISN  420
EYKGMPGGRN LYKIQLIDYT KVTCLSCELN PERCQYYSVS FSKEAKYYQL RCSGPGLPLY  480
TLHSSVNDKG LRVLEDNSAL DKMLQNVQMP SKKLDFIILN ETKFWYQMIL PPHFDKSKKY  540
PLLLDVYAGP CSQKADTVPR LNWATYLAST ENIIVASFDG RGSGYQGDKI MHAINRRLGT  600
FEVEDQIEAA RQFSKMGFVD NKRIAINGWS YGGYVTSMVL GSGSGVFKCG IAVAPVSRWE  660
YYDSVYTERY MGLPTPEDNL DHYRNSTVMS RAENFKQVEY LLIHGTADDN VHFQQSAQIS  720
KALVDVGVDF QAMWYTDEDH GIASSTAHQH IYTHMSHFIK QCFSLP
```

```
1.         11         21         31         41         51
|          |          |          |          |          |
ATGGCTCTGG CGAAGGTGAG GGAGCCAAAC GCAAATGACA ATGCCATCAG AGTTGACAAC   60
AGAAGTGTGA TTAAAGTGCG TGCTAACCAG TGTTCCCTGC ATGAGGCAGA AAGTGAATCC  120
AGAAACCCTC AGGAGCTCTG GATGGGCCTG CTCCTCTTGA TGGGGGTCCT AGAAGCATGT  180
GTGGAAATGA GGCCTCTGTC AGTCTGGTCC CTGAGAGATG ACAAGGAGCA GAGCCCCCAC  240
CAGCCCACAC TGGATGTCTA A
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MALAKVRERN ANDNAIRVDN RSVIKVRANQ CSLHEAESES RNPQELWMGL LLLMGVLEAC   60
VEMRPLSVWS LRDDKEQSPH QPTLDV
```

A207 DNA SEQUENCE
Gene name: ESTs
Unigene number: Hs.222886
Probeset Accession #:   AI672225
Nucleic Acid Accession #:   none found
Coding sequence: 1-462 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCCAAATG CTGAGTTAGA AGCAAAGAGC CTTGGAAGCA GTAAATGTTT AAAAACTGCT    60
CTCATACTTG CTGTATGTTG TGGATCAGCA AATATAGTCA GCCCTCTACT TGAGCAAAAT   120
ATTGATGTAT CTTCTCAAGA TCTGGACAGA CGGCCAGAGA GTATGCTGTT TCTAGTCATC   180
ATCATGTGGA CCAGTTTTGT GGAAGACAAT CTTTCCATGG GCTGGGGGAA GCTAGAAGAT   240
TTTATGGCTA TTGAAGAAGA AATGAAGAAG CACGGAAGTA CTCATGTGGG ATTCCCAGAA   300
AACCTGACTA ATGGTGCCGC TGCTGGCAAT GGTGATGATG GATTAATTCC TCCAAGGAAG   360
AGCAGAACAC CTGAAAGCCA GCAATTTCCT GACACTGAGA ATGAAGAGTA TCACAGGTTT   420
GTCAAAGATC AGATAGTTGT AGATATGCGG CGTTATTTCT GA                      
```

A208 Protein sequence:
Gene name: ESts
Unigene number: Hs.222886
Probeset Accession #:   AI672225
Protein Accession #:   none found
Signal sequence: none found
Transmembrane domains: 16-38
Cellular Localization: not determined

```
1          11         21         31         41         51
|          |          |          |          |          |
MPNAELEAKS LGSSKCLKTA LILAVCCGSA NIVSPLLEQN IDVSSQDLDR RPESMLFLVI    60
IMWTSFVEDN LSMGWGKLED FMAIEEEMKK HGSTHVGFPE NLTNGAAAGN GDDGLIPPRK   120
SRTPESQQFP DTENEEYHRF VKDQIVVDMR RYF
```

A209 DNA SEQUENCE
Gene name: odz (odd Oz/ten-m, Drosophila) homolog 1
Unigene number: Hs.23796
Probeset Accession #:   NM_014253
Nucleic Acid Accession #:   NM_014253
Coding sequence: 65-8242 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
GACTGCTTGC ATTAAAGGAC TTCCTCATCC TTTTTTTCAT GAAACTGAGC TTGCTTAATC    60
AGAGATGGAG CAAACTGACT GCAAACCCTA CCAGCCTCTA CCAAAAGTCA AGCATGAAAT   120
GGATCTAGCT TACACCAGTT CTTCTGATGA GAGTGAAGAT GGAAGAAAAC CAAGACAGTC   180
ATACAACTCC AGGGAGACCC TGCACGGAGTA TAACCAGGAG CTGAGGATGA ATTACAATAG   240
CCAGAGTAGA AAGAGGAAAG AAGTAGAAAA ATCTACTCAA GAGATGGAAT TCTGTGAAAC   300
CTCTCACACT CTGTGCTCTG GCTACCAAAC AGACATGCAC AGCGTTTCTC GGCATGGCTA   360
CCAGCTAGAG ATGGGATCTG ATGTGGACAC AGAGACAGAA GGTGCTGCCT CACCTGACCA   420
TGCACTAAGA ATGTGGATAA GGGGAATGAA ATCAGACGAT AGTTCCTGTT TGTCCAGCCG   480
GGCCAACTCT GCATTATCCT TGACTGACAC TGACCATGAA AGGAAGTCTG ATGGGGAAAA   540
TGGTTTCAAA TTCTCTCCTG TTTGTTGTGA CATGGGAGGCT CAAGCTGGGT CTACTCAAGA   600
TGTGCAGAGC AGCCCACACA ACCAGTTCAC CTTCAGACCC CTCCCACCGC CACCTCCGCC   660
TCCTCATGCC TGCACCTGTG CCAGGAAGCC ACCCCCTGCA GCGGACTCTC TTCAGAGGAG   720
ATCAATGACT ACCCGCAGCC AGCCCAGCCC AGCTGCTCCA GCTCCCCCAA CCAGCACGCA   780
GGATTCAGTC CATCTGCATA ACAGCTGGGT CCTGAACAGC AACATACCAT TGGAGACCAG   840
GCATTCCCTG TTCAAACATG GATCTGGTTC CTCTGCGATC TTCAGTGCAG CCAGTCAGAA   900
CTACCCTCTG ACATCCAATA CCGTGTACTC GCCCCCTCCC AGGCCTCTTC CTCGAAGCAC   960
CTTTTCCCGA CCTGCCTTTA CCTTTAACAA ACCTTACAGG TGCTGCAACT GGAAGTGCAC  1020
AGCATTGAGC GCCACTGCAA TCACAGTGAC TTTGGCCTTG TTACTAGCCT ATGTGATTGC  1080
AGTGCATTTG TTCGGCCTGA CTTGGCAGTT GCAACCAGTT GAAGGAGAGC TGTATGCAAA  1140
TGGAGTTAGC AAAGGGAACA GGGGGACCGA GTCCATGGAC ACTACTTACT CTCCAATTGG  1200
AGGAAAAGTT TCTGATAAAT CAGAGAAAAA AGTGTTTCAG AAGGGACGGG CGATAGACAC  1260
TGGAGAAGTT GACATTGGTG CACAGGTCAT GCAGACCATT CCACCTGGTT TATTCTGGCG  1320
TTTCCAGATT ACTATCCACC ATCCAATATA TCTGAAGTTC AATATTTCTT TAGCCAAGGA  1380
CTCTCTGCTG GGAATTTATG GCAGAAGAAA CATTCCACCT ACACATACTC AGTTTGATTT  1440
TGTAAAACTA ATGGATGGCA AACAGCTGGT CAAGCAGGAC TCCAAGGGCT CTGATGAATAC  1500
ACAGCACTCC CCTCGGAACC TGATCTTAAC TTCGCTTCAG GAGACAGGTT TCATAGAGTA  1560
TATGGATCAA GGACCTTGGT ATCTGGCGTT TTACAATGAT GGAAAAAAGA TGGAGCAAGT  1620
ATTCGTGTTA ACTACAGCAA TTGAAATAAT GGATGACTGT TCAACCAATT GCAATGGAAA  1680
TGGAGAGTGT ATCTCTGGCC ATTGTCATTG TTTCCCAGGA TTCCTTGGAC CTGACTGTGC  1740
TAGAGATTCC TGCCCTGTGC TGTGTGGTGG GAATGGAGAA TACGAGAAAG GACACTGTGT  1800
CTGCCGGCAT GGCTGGAAGG GGCCAGAGTG TGACGTTCCG GAAGAACAAT GCATTGATCC  1860
AACATGCTTT GGCCACGGCA CCTGCATCAT GGGAGTCTGC ATCTGTGTGC CAGGATACAA  1920
AGGAGAAATA TGCGAGGAAG AGGACTGCCT AGACCCAATG TGTTCCAACC ATGGCATCTG  1980
TGTAAAGGGA GAATGTCACT GTTCCTACTGG CTGGGGAGGA GTTAACTGTG AAACACCACT  2040
TCCTGTATGT CAAGAGCAGT GCTCAGGACA CGGAACTTTT CTTCTGGACG CTGGAGTATG  2100
CAGCTGTGAT CCCAAGTGGA CAGGATCTGA CTGCTCAACA GAGCTGTGTA CCATGGAGTG  2160
TGGTAGCCAT GGAGTCTGCT CAAGAGGAAT TTGCCAGTGT GAAGAAGGCT GGGTAGGACC  2220
AACATGTGAG GAACGCTCCT GTCATTCTCA TTGTACTGAG CATGGCCAAT GCAAAGATGG  2280
```

```
AAAATGTGAG TGTAGCCCTG GATGGGAGGG CGACCACTGC ACAATTGCTC ACTACTTAGA 2340
TGCTGTCCGA GATGGCTGCC CAGGGCTCTG CTTTGGAAAT GGACGATGTA CCCTGGATCA 2400
AAATGGTTGG CACTGTGTGT GTCAGGTGGG TTGGAGTGGG ACAGGCTGCA ATGTTGTCAT 2460
GGAAATGCTT TGTGGAGATA ACTTGGACAA TGATGGAGAT GGTTAACCG ACTGTGTGGA 2520
TCCTGACTGT TGTCAACAAA GCAACTGTTA TATAAGTCCT CTCTGCCAGG GCTCACCAGA 2580
TCCTCTTGAC CTCATTCAGC AAAGCCAAAC TCTCTTCTCT CAGCACACTT CAAGACTTTT 2640
TTATGATCGA ATCAAATTCC TCATTGGCAA GGACAGTACT CATGTCATTC CTCCTGAGGT 2700
GTCATTTGAC AGCAGGCGTG CCTGTGTGAT TCGAGGCCAA TGGTGGCCA TAGATGGAAC 2760
TCCTCTAGTG GGAGTGAATG TCAGTTTCTT GCACCACAGT GATTATGGGT TTACCATCAG 2820
CCGGCAAGAT GGAAGCTTTG ACCTCGTGGC CATCGGTGGC ATCTCTGTCA TCTTAATCTT 2880
CGACCGATCC CCTTTCCTGC CTGAGAAGAG AACACTCTGG TTGCCTTGGA ATCAGTTTAT 2940
TGTGGTAGAG AAAGTCACCA TGCAGAGAGT TGTATCGAGC CCGCCATCCT GCGATATCTC 3000
CAACTTTATC AGCCCAAACC CTATTGTGCT TCCTTCACCG CTCACATCAT TTGGAGGGTC 3060
CTGTCCAGAG AGGGGAACTA TTGTTCCTGA GCTGCAGGTT GTACAGGAGG AAATTCCCAT 3120
TCCCTCCAGC TTTGTGAGGC TGAGTTACCT GAGCAGCCGC ACCCCTGGGT ATAAAACCCT 3180
GCTACGGATC CTTCTGACAC ATTCAACGAT TCCCGTAGGC ATGATAAAAG TACACCTCAC 3240
AGTAGCTGTG GAAGGGCGAC TCACACAGAA GTGGTTCCC GCCGCAATTA ATCTTGTCTA 3300
CACATTTGCT TGGAACAAGA CCGATATCTA TGGACAGAAG GTTTGGGGCC TGGCAGAGGC 3360
TTTGGTATCT GTGGGATATG AATATGAAAC GTGCCCTGAC TTTATTCTCT GGGAGCAAAG 3420
GACAGTCGTT TTACAAGGTT TTGAGATGGA TGCTTCTAAC CTAGGAGACT GGTCTTTGAA 3480
TAAGCATCAC ATTTTGAATC CTCAAAGTGG AATCATACAT AAAGGGAATG GAGAAAAATAT 3540
GTTCATTTCC CAGCAGCCCC CAGTCATATC AACCATAATG GGTAATGGAC ACCAAAGGAG 3600
TGTAGCCTGC ACCAACTGCA ATGGCCCAGC CCACAACAAC AAACTCTTTG CTCCTGTCGC 3660
CTTAGCTTCT GGCCCTGATG GCAGTGTGTA TGTTGGCGAC TTCAATTTTG TAAGGAGAAT 3720
ATTTCCCTCG GGAAACTCCG TTAGTATTTT GGAATTAAGC ACAAGTCCTG CTCACAAATA 3780
CTATCTGGCT ATGGACCCTG TGTCTGAATC ACTCTATCTA TCAGACACCA ATACTCGCAA 3840
AGTCTACAAG TTGAAATCTC TTGTGGAGAC GAAAGATCTG TCCAAGAATT TTGAAGTGGT 3900
GGCAGGAATT GGTGATCAGT GCCTTCCCTT TGACCAGAGT CATTGTGGAG ATGGTGGGAG 3960
AGCATCGGAA GCTTCACTGA ATAGCCCTCG AGGCATCACA GTTGATAGGC ATGGATTTAT 4020
TTACTTTGTG GATGGGACTA TGATTCGCAA AATTGATGAG AATGCTGTGA TCACAACTGT 4080
AATCGGCTCA AATGGTCTGA CTTCCACACA ACCACTGAGC TGTGACTCAG GAATGGACAT 4140
CACTCAGGTG CGATTAGAGT GGCCAACAGA CCTTGCAGTA AATCCTATGG ACAATTCATT 4200
GTATGTCTTG GATAACAACA TTGTGCTGCA AATTTCTGAG AACAGGCGTG TTCGGATCAT 4260
CGCAGGACGC CCCATTCACT GCCAGGTGCC AGGCATCGAT CATTTCCTGG TCAGCAAGGT 4320
AGCAATTCAC TCCACTCTAG AGTCAGCGAG GGCCATCAGT GTCTCCCACA GCGGGCTGCT 4380
CTTCATAGCT GAAACAGACG AGAGGGAAAGT AAACCGCATT CAGCAAGTAA CCACCAATGG 4440
GGAGATCTAC ATCATCGGCTG GTGCCCCCAC TGACTGTGAC TGCAAAATTG ATCCAAACTG 4500
TGACTGTTTT TCAGGTGATG GTGGCTATGC CAAAGATGCA AAGATGAAAG CCCCCTTCCTC 4560
CTTAGCAGTG TCGCCTGATG GAACCCTCTA TGTGGCAGAC CTCGGAAATG TTCGAATTCG 4620
TACCATCAGC AGGAACCAAG CCCACCTGAA TGACATGAAC AITTATGAGA TTGCTTCACC 4680
CGCTGATCAG GAACTGTACC AGTTCACTGT AAATGGAACC CACCTACACA CCCTGAACTT 4740
GATAACAAGG GACTATGTTT ATAACTTCAC CTACAATTCT GAAGGTGACT TGGGCGCGAT 4800
TACCAGCAGC AATGGCAATT CAGTGCACAT TCGCCGTGAT GCAGGCGGAA TGCCGCTATG 4860
GCTTGTGGTG CCTGGCGGAC AAGTATACTG GCTGACTATA AGCAGCAATG GAGTCCTGAA 4920
AAGAGTGTCA GCCCAAGGCT ATAATCCGGC CTTAATGACC TATCCAGGAA ACACAGGGCT 4980
TCTGGCTACC AAAAGTAACG AAAATGGGAT GACAACGGTT TATGAGTATG ACCCCGAGGG 5040
ACACCTGACC AATGCAACGT TTCCCACTGG AGAGGTCAGC AGCTTCCACA GTGACCTGGA 5100
GAAGCTGACA AAAGTGGAGC TAGATACTTC CAACCGTGAA AATGTCCTCA TGTCAACCAA 5160
CTTGACGGCA ACTAGTACCA TATATATTTT AAAACAAGAA AATACTCAAA GTACCTATCG 5220
GGTGAATCCA GATGGTTCCC TGCGTGTCAC TTTTGCCAGC GGGATGGAGA TCGGCCTCAG 5280
CTCAGAGCCC CACATCCTGG CAGGGGCAGT CAACCCTACC CTGGGCAAAT GCAACATCTC 5340
ATTGCCCGGA GAGCACAATG CAAACCTCAT CGAGTGGCGG CAGAGGAAGG AGCAAAACAA 5400
AGGCAATGTT TCGGCTTTTG AAAGGAGGCT GAGGGCCCAC AACAGAAACC TACTCTCCAT 5460
AGATTTTGAT CATATAACCC GCACAGGAAA GATCTATGAT GACCATCGAA AATTCACCCT 5520
TCGAATTCTT TATGACCAGA CTGGGCGACC CATTCTGTGG TCTCCTGTAA GCAGATATAA 5580
TGAAGTGAAC ATCACATATT CACCTTCGGG ATTGGTGACG TTTATTCAAA GAGGAACGTG 5640
GAATGAAAAA ATGGAATATG ACCAGAGTGG GAAAATTATT TCAAGAACTT GGGCTGATGG 5700
GAAAATTTGG AGCTATACCT ACTTAGAAAA ATCTGTGATG CTTCTCCTAC ACAGCCAGCG 5760
GCGTTACATC TTTGAGTATG ACCAATCAGA TTGCCTGCTG TTACCTACCA TGCCTAGCAT 5820
GGTGCGCCAC AGCTTACAAA CCATGCTTTC AGTGGGCTAC TACCGTAATA TCTACACCCC 5880
ACCGACAGT AGCACTTCTT TTATCCAAGA CTATAGTCGA GATGGCCGAT TGCTACAGAC 5940
CCTGCATCTG GGGACAGGGC GCAGAGTCTT ATACAAGTAC ACCAAGCAAG CAAGGCTTTC 6000
TGAGGTTCTC TATGATACCA CTCAGGTCAC ATTAACATAT GAAGAGTCTT CTGGAGTGAT 6060
TAAGACAATA CACCTGATGC ATGACGGATT CATCTGCACA ATCAGATACA GGCAAACAGG 6120
ACCTCTTATT GGACGCCAGA TTTTCAGATT CAGTGAAGAA GGCCTTGTGA ATGCACGGTT 6180
CGACTACAGC TACAACAATT TCCGAGTCAC AAGCATGCAA GCTGTAATCA ATGAAACCCC 6240
TTTGCCTATA GATCTTTACC GATATGTTGA TGTCTCTGGC AGAACAGAGC AGTTTGGAAA 6300
ATTCAGTGTA ATTAATTACG ATTTAAATCA GGTCATAACT ACTACAGTGA TGAAACACAC 6360
CAAAATCTTC AGTGCCAATG GACAAGTCAT TGAAGTCCAA TATGAAATCC TAAAGGCAAT 6420
TGCCTACTGG ATGACCATTC AATATGATAA TGTGGGCCGA CATGGTAATA TGTGCATAAG 6480
GGTAGGAGTA GATGCCAATA TAACAAGGTA CTTCTATGGA TACGATGCTG ATGGGCAACT 6540
TCAGACTGTT TCTGTAAATG ACAAAACCCA GTGGCGTTAT AGTTACGATC TGAATGGGAA 6600
CATCAACCTC TTAAGCCATG GGAAGAGTGC TGGTCTTACT CCTCTCCGAT ATGACCTCCG 6660
AGACCGCATC ACCAGATTAG GAGAAATTCA GTATAAAATG GATGAAGATG GCTTTCTGAG 6720
GCAGAGGGGA AATGATATTT TTGAAATATAA TTCTAATGGC CTGCTGCAGA AAGCCTACAA 6780
TAAGGCTTCT GGCTGGACTG TGCAGTATTA CTATGATGGG CTTGGGCGAC GTGTCGCGAG 6840
TAAGTCCAGC CTAGGGCAGC ACCTTCAGTT CTTTGTCGAC GCGACCGCGA ACCCCATAAG 6900
AGTTACTCAT TTGTACAACC ACACAAGCTC GGAGATTACA TCTCTGTATT ATGATCTCCA 6960
AGGTCACCTT ATTGCCATGG AGTTAAGCAG TGGTGAAGGA TATTATGTAG CCTGTGATAA 7020
TACAGGTACC CCACTAGCTG TGTTCAGCAG CCGAGGTCAG GTCATAAAGG AGATACTATA 7080
CACACCTTAT GGCGATATCT ATCATGACAC TTACCCTGAC TTTCAGGTCA TAATTGGTTT 7140
TCATGGAGGA CTCTATGATT TCCTTACTAA ATTAGTGCAC CTGGGGCAAA GGGATTATGA 7200
TGTTGTTGCT GGCAGATGGA CAACGGCCTA TCATCACATA TGGAAACAGT TGAACCTCCT 7260
TCCTAAACCA TTCAACCTCT ACTCCTTTGA AAATAACTAC CCAGTTGGCA AAATTCAAGA 7320
```

```
        TGTTGCAAAG TATACCACAG ACATCAGAAG TTGGTTGGAG CTATTTGGTT TCCAATTACA  7380
        CAATGTACTA CCTGGATTTC CCAAACCTGA ATTAGAAAAT TTAGAATTAA CTTACGAGCT  7440
        TCTACGGCTT CAGACAAAAA CTCAAGAGTG GGATCCTGGA AAGACTATCC TGGGCATTCA  7500
        GTGTGAACTC CAGAAACAGC TCAGGAATTT CATTTCCTTG GACCAACTAC CTATGACTCC  7560
        CCGATACAAT GATGGACGGT GCCTTGAAGG AGGGAAGCAA CCAAGGTTTG CTGCTGTCCC  7620
        TTCTGTTTTT GGGAAAGGTA TAAAATTTGC CATCAAGGAT GGCATAGTAA CAGCTGATAT  7680
        TATAGGAGTA GCCAATGAAG ATAGCAGGCG GCTTGCTGCC ATTCTCAATA ATGCCCATTA  7740
        CCTGGAAAAC CTACATTTTA CCATAGAGGG GAGGGACACT CACTACTTCA TTAAGCTTGG  7800
        GTCTCTGGAG GAAGACCTGG TGCTCATCGG TAACACTGGG GGGAGGCGGA TTCTGGAGAA  7860
        TGGTGTCAAT GTCACTGTGT CCCAGATGAC TTCTCTGTTG AATGGGAGGA CTAGACGGTT  7920
        TGCAGATATT CAGCTCCAGC ATGGAGCCCT GTGCTTCAAC ATCCGGTATG GGACAACTGT  7980
        CGAAGAGGAA AAGAATCACG TGTTGGAGAT TGCCAGACAG CGCGCAGTGG CCCAGGCCTG  8040
        GACTAAGGAA CAAAGAAGGC TGCAAGAGGG GGAAGAGGGG ATTAGGGCAT GGACAGAAGG  8100
        GGAAAAGCAG CAGCTTTTGA GCACTGGGCG GGTACAAGGT TACGATGGGT ATTTTGTTTT  8160
        GTCTGTTGAG CAGTATTTAG AACTTTCTGA CAGTGCCAAT AATATTCACT TTATGAGACA  8220
        GAGCGAAATA GGCAGGAGGT AACAAAAATA TCTCTGCCTT TGCGTCACCA AAGACTGCCT  8280
        GTTTTTAAAA CATAAAATGG TTTATTGTAT TGGTTTTCTA GATCAGAACT CTGTATATGT  8340
        AAATATGGAG GAAAAACATA TCCAACTGCC TTTCAATGTG ACGGAAGATG GTATTTTAAT  8400
        ATTGTTTGTT TAAACTCTTT AAGAAATGAC AGAGATTTTT AGTTCTTGTG TGGCAGTATT  8460
        CAAAATAACA CAAGTAGAAC TCAAACAGCT AAAAACAGTT TTCAGAAAGC ACCACTTTCA  8520
        ATTTGCCGAG CCATGCATAT GTTCCAATAT CCAGAAAGAA CCCAAGGTTC TCTATCTCTA  8580
        TTGTGAGAAG CAGTTTCATC CTTAACTGTT GGCAGAACTT ACGGGCTATT TGAATAGGTG  8640
        GTGCAATAGT ATCTGAAACT TGCCTTTCGA AAGACTGCCA GCCCTTTGAC GTTTTCCAGA  8700
        TCTGTTATAG GAAACTTAAA AACAGGTGTA AAATGTCTTC AGCCACCATC TCCTAGAGTG  8760
        AGGACCCAAT TGCCCTTCCT TCTTGATTAT TCCTCCTTGC TTGTTAAAGT AAATGCCATA  8820
        TTGTTGTGCT GTGTTTTGGC GTGTGGTGGC TGGGTTCTGT CTACCATGCT TCCCTGTGGG  8880
        TGTGGTAACC AGACTGTATA GCCGCTATTT GCTCGTGTGT ACATGATACC AAAGCAGCTG  8940
        GCCAGCGTGA CCTCTCTCAC ACGACCTGTT TTGACTCAAT TTTTTACTAA AAGTTGTTCA  9000
        GCTGTATTGG TATCATGTAA ACATAGCTTT TATTAACCTG GGTAGGAATT TCTCATTTAT  9060
        ATATAGGATG TGTTTTGGTC ATAGTTTCAC ATTAGTGATT CAGTATCTAT ACACTGACCC  9120
        AATGGTTTTG TGCACATGAA CGGTAATTTA CTTAAAAGTA TGATTCTGGT ACAAAAACAA  9180
        ACAAAGGCTT TAGCAGGCAT ACGTGTCTGG GATGCCGATA CATACATTAA CTACTACTGC  9240
        AGAAATTCAT AAGAGCCAAA ACCTTAAAAA AATAGACATG GTACTTAAGT GAAAGTACTA  9300
        AAGGGAAGAC CAGACCCAAAC ATCACAGCAG TTGCTGCCAC ATTGTTTCAG CCCACTTAGA  9360
        TTTATCTTTC AAATGTACAA TTCTGTATTG AACATCTCCC AGCCATCTTC AGGAAATCGA  9420
        ATCAAGTAAA TCCTTTCCAA CCGAAAACAT TTCAACTAAC TATAGAGAGG CAGACTGACT  9480
        TTTACTAAAA TAATTTATAC AGTTAGTTAT TTCGTTCTC CGTACTTACC CATTTATCTT  9540
        TATTTAATCG TCTCTACTGC CTAGGAAAAT AACTATTTTC CAGGACGGGT TATTTGTTCT  9600
        GCGATCATTT AAAATTTGGA GAAAGGTCAG GATTAGTCTAT AATATCAGCT GCAGTTTCTC  9660
        AATCTCTAGG AATCCTGCAG TAAAACAAGC CCCTTGGTGA GCTGGAAGAT TTGTGCCCAG  9720
        TGACAAAGAG ATAGTTTGTA AAATGCTGTG TAATTGTAAG TTACCACAAA TGAAAATACA  9780
        TGACAGCACA ATGTGGCCCG TAGAAAATTC CCCTGAGCCA GCTTCTGCAC TTTCATCACG  9840
        GAATCTGAAC ATTTGCTATG TCTGAAGGCA AATTTATGAT GGAATGTTAG TTTGGATTCT  9900
        TTCCAGATGC TACCTAAATG CAGTGTGGGG TCATTGCCTT GCTTTGCGAT GACAGTTTCT  9960
        TTGAAAAATAT GCAAAGTCAT AAGCTCATGT TAAGGTTTTT CAAGAGTCTG CCTCCTACTA  10020
        CACAAAGGAA AGCAAGGGAA AGGAAATGAC CCTGGCAAAC AGTAGGGAAG GGTGTATTCA  10080
        AACATTTCAT TTTCAAAACC TTCGGGGTTAG AATACCACTT ACACATGTAT TCTGAGAGAC  10140
        AGAATTCATG AGGAACTCAT CTCTCTTTAT AACTGGAAAC ACACCAGCTT GATATATTGC  10200
        TAATCCATAC TAAAATCATA TTATTGGGTT TTTTCTGAAT CAGGCCTGTA TTAATGGTAC  10260
        AGTATTTATT CAGAATGGAA TTCTAAAATT ACTAACAAAC TTGTTGAAAA TTTGAATACC  10320
        TCCACACCAA CCTAAAAATG GACCTTAAGT TCCTAGAACC TCTGATGTTC TTTTAAATTA  10380
        ATGGAAAAAT AATTGTGAA CTGTATATAG AGAGTGCATT CATAAATGTG ATTATGTATT  10440
        TTATCACAAA TCCAAAATGT CAATATTAGA GTCTATTTTG CTTATATTTT AAGCAATTAT  10500
        ACGTTTTTGC AATTCATTGA TGATGTATCA TTTTCAAACT GCTTTAAAATA TCCATTAGAA  10560
        ACAAATATTT GAAGCTTTTA CTTAAATAGTG ATTACCTTGA ACTGTGCATT TCTAGTTTGT  10620
        AATACGTATT TGGTTGGTTC GTGCCTTTAG TTTGTTAAAG TTACATTTGT ATTATATTCA  10680
        GGAAATGCAC TTTTTATTAC TTACAGCTGT GGTTTTAATA CTGCCTTGAA CTATTATTAT  10740
        TCTTTTTACA ACTCCTAAAG CTTGAGGGAG GAAAGAAAAA AAAAACAAAA CTACTAATCA  10800
        GTAGTAAATC GAAGAGAAAC ATTTTGGCAT TTCTTAAGAA GAAGATGGAG ATATTGAGTA  10860
        TATCACTTCC TATTCAGCTG AATAGAAAGA ATGCCTTCAT TGACTTGCAG TTCTGCAGTT  10920
        TAAATTATTG AAAGAACAAT TCGTTTGCAT TTCCTGATGA AAGTAAAAGC ATTTTTCAGA  10980
        GAAACATATG AATTTCTCAT ACCCAGCAGA CAGATGGCTG ACACTGCACA GCCACACACC  11040
        ATTCGAGTAA GTTAAAGTGA GAGCATAGTA GTTGGACTCT CCTATGAAGA ACATTCTGGG  11100
        CTGGAGGCAG GGAATACTCC ATGGTTGTTT CTTTTTCCTA CTTAAGCCCA TTTTGTTTGT  11160
        GCTTTTCTGT TTTGTTTTGT TTTCACTCTT GCACTACAGT CTAGAGATCC AAATGAACTG  11220
        AAAAGTTCAA AGTTTAACAC ATTTAAATAT GTTTACTTTT AGTTGTCATT CTAATCGTTA  11280
        TTGATTAGAA GCATGACTCC TGAAGGAAAG GGAAATAAAT CTCAATTCAT ACTAACTTGC  11340
        AACAAAACAC TTTTACCATA TAAATAAGTA TATGATTTAT TTTTAACCCA AAAAATGTAT  11400
        AAAATAAGTG TGTCCTTTAC TGTCAATTTA TCGAGAAGAT CTATAATATA TAGACTACAT  11460
        ATATATAATA TATACAACAT AGCCAAATGT ATGAAAACTT GACAATGTAT AATTTGGAAT  11520
        TCACATGCTA CCTATGTAGA CAGGTATGAA ATTAAGTTAT AATTTTCATG AGACATTTTC  11580
        ATCACTGTTG ACACAGTTTC AAGGCATTCC ATCATGTTAT TTTGACTCTT TTTCTTTTTT  11640
        TTTTCTTTAA AAATATATTT TTAACTAGAC CAGGCCCCAC TATAATATCA CTTAAGAGAG  11700
        TCAGGGCAAA GTTTTTGCAT TTATGAAGAT GTGTTCATGT AAGGGTGATT GTAATGGAGT  11760
        TCATTGGTAA TAGAAGCAAA AGTACAGTAA CGAAGTATTG AAAAGAAAAT TTGGAGACA  11820
        TTGGAGCATA TTATATATAG CTTGTGGAAA GACATAAGGC TACAGATGGA ATGGAACATT  11880
        CCTGTTTTCT TGAAGAAATT CACATACACA TAGCTGACCT GACTAGTGCT TCAGCTCTTC  11940
        CACAGCCTTC TATAAAGGTT CTTTCTTCTG CAAAGAAAAC AAAACAAAAC AAAACAAAAC  12000
        AAAAAAAAAC AAAAAAAGCG CAAAAAACAA AAAAACAAAA AAAAGCAAAG TAAAATTTAA  12060
        AAATACAGAA AACAAACAAC AAAAAAGAAT TCAACCATAA ATAGTGACTA TTATTTTCAG  12120
        TGTGTCCTTC ATGTGAAAGC TATTAAGGAC CAAATATACT ACTGTTCATA AGAAGAAATT  12180
        ACTTTCTAAA CAGTAACTGA AAATACTTAG AGTTAAACTT GCTGTGGATT TTGTCTTGGC  12240
        AGTTGTCATC TTACATTATT TGTCAAAGGA AATGTGTTTG GCAGTTAAAA ATCTTTCCTT  12300
        AGATTTAGTG GTGGACTTTA ACCTCTTAAA TAAATGTTAG TATATCAGAT TGTGTCCTTG  12360
```

930

```
TCACCAGATT TGGTTCTTAA ACTTT
```

Seq ID NO: 272 <u>Protein sequence</u>
Protein Accession #: NP_006465.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MWKVSALLFV LGSASLWVLA EGASTGQPED DTSTTGLEGG VAMPGAEDDV VTPGTSEDRY   60
KSGLTTLVAT SVNSVTGIRI EDLPTSESTV HAQEQSPSAT ASNVATSHST EKVDGDTQTT   120
VEKDGLSTVT LVGIIVGVLL AIGFIGGIIV VVMRKMSGRY SP
```

Seq ID NO: 273 <u>DNA sequence</u>
Nucleic Acid Accession #: CAT cluster

```
1          11         21         31         41         51
|          |          |          |          |          |
GCGGCCGCCA GCTTGCAAAG CCGAAGTCTG GCCGCGCTCT TCGACTCGCT GCGCCACGTC   60
CCCGGGGGTG CCGAGCCGGC GGGGGGTGAG GTGGCTGCGC CGGCGGCCGG GCTAGGAGGT   120
GCGGGCACTG GGGGCGCGGG AGGGGACGTG GCAGGCCCCG CGGGGGCCAC GGCGGATCCCA   180
GGGGCCAGGA AGGTCCCGCT GCGGGCACGC AATCTGCCTC CGTCCTTCTT CACGGAGCCG   240
TCCGGGGCAG GCGGGCGGCGG GTGTGGCCCG TCGGGGCCGG ACGTGAGCTT GGGCGGACCTG   300
GAGAAGGGCG CGGAGGCCGT GGAGTTCTTT GAGCTGCTGG GGCCGGACTA CGGCGGCCGGC   360
ACGGAGGCGG CAGTCTTGCT TGCCGGCCGAG CCTCTCGACG TGTTCCCCGC CGGAGCCTCC   420
GTACTGCGGG GACCCCCGGA GCTGGAGCCC GGGCCTCTTTG AGCGCGGCCGCC GGCAGTGGTG   480
GGAAACCTAC TGTACCCCGA GCCCTGGAGC GTCCCGGGCT GCTCCCCGAC CAAAAAGAGC   540
CCCTTGACTG CCCTCCGCGG CGGGGTTGACC TTGAACGAGC CCTTGAGCCC CCTGTACCCC   600
GCCGCTGCGA ATTTCTCCCG GCGGGGAGGA CGGGCCGGGC CATTTGGCTT CTTTCGCCCC   660
CTTCTTTCCA GACTGCGCTT TGC
```

Seq ID NO: 274 <u>DNA sequence</u>
Nucleic Acid Accession #: Eos sequence

```
1          11         21         31         41         51
|          |          |          |          |          |
CAAAGAGGCC GGGCTCCAGC TCCGGGGGTC CCCGCAGTAC GGAGGCTCCG GCGGGGAACA   60
CGTCGAGAGG CTCGGCGGCA AGCAAGACTG CCGCCTCCGT GCCGGCCGCCG TAGTCGGGCC   120
CCAGCAGCTC AAAGAACTCC ACGGCCTCCG CGCCCTTCTC CAGGTCGCCC AAGCTCACGT   180
CCGGCCCCGA CGGGCCACAC CCGCCGCCGC CTGCCCGGGA CGGCTCCGTG AAGAAGGACG   240
GAGGCAGATT GCGTGCCCGC AGCGGGACCT TCCTGGCCCC TGGGATCGCC GTGGCCCCCG   300
CGGGGCCTGC CACGTCCCCT CCCGCGCCCC CAGTGCCCGC ACCTCCTAGC CCGGCCGCCG   360
GCGCAGCCAC CTCACCCCCC GCCGGCTCGG CACCCCCGGG GACGTGGCGC AGCGAGTCGA   420
AGAGCGCGGC CAGACTTCGG CTTTGCAAGC TGGCGGCCGC
```

Seq ID NO: 275 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_001118.1
Coding sequence: 74..1651

```
1          11         21         31         41         51
|          |          |          |          |          |
AGCCCAGAGA CACATTGGGG CTGACCTGCC GCTGCTGTCA GTGGGAGGCC AGTGGTGCTG   60
GCCAAGAAGT GTCATGGCTG GTGTCGTGCA CGTTTCCCTG GCTGCTCACT GCGGGGCCTG   120
TCCGTGGGGC CGGGGCAGAC TCCGCAAAGG ACGCGCAGAC TGCAAGTCCG CGGCCCAGAG   180
ACACATTGGG GCTGACCTGC CGCTGCTGTC AGTGGGAGCC CAGTGGTGCT GGCCAAGAAG   240
TGTCATGGCT GGTGTCGTGC ACGTTTCCCT GGCTGCTCTC CTCCTGCTGC CTATGGCCCC   300
TGCCATGCAT TCTGACTGCA TCTTCAAGAA GGAGCAAGCC ATGTGCCTGG AGAAGATCCA   360
GAGGGCCAAT GAGCTGATGG GCTTCAATGA TTCCTCTCCA GGCTGTCCTG GGATGTGGGA   420
CAACATCACG TGTTGGAAGC CCGCCCATGT GGGTGAGATG GTCCTGGTCA GCTGCCCTGA   480
GCTCTTCCGA ATCTTCAACC CAGACCAAGT CTGGGAGACC GAAACCATTG GAGAGTCTGA   540
TTTTGGTGAC AGTAACTCCT TAGATCTCTC AGACATGGGA GTGGTGAGCC GGAACTGCAC   600
GGAGGATGGC TGGTCGGAAC CCTTCCCTCA TTACTTTGAT GCCTGTGGGT TTGATGAATA   660
TGAATCTGAG ACTGGGGACC AGGATTATTA CTACCTGTCA GTGAAGGCCC TCTACACGGT   720
TGGCTACAGC ACATCCCTCG TCACCCTCAC CACTGCCATG GTCATCCTTT GTCGGTTCCG   780
GAAGCTGCAC TGCACACGCA ACTTCATCCA CATGAACCTG TTTGTGTCGT TCATGCTGAG   840
GGCGATCTCC GTCTTCATCA AAGACTGGAT TCTGTATGCG GAGCAGGACA GCAACCACTG   900
CTTCATCTCC ACTGTGGAAT GTAAGGCCGT CATGGTTTTC TTCCACTACT GTGTTGTGTC   960
CAACTACTTC TGGCTGTTCA TCGAGGGCCT GTACCTCTTC ACTCTGCTGG TGGAGACCTT   1020
CTTCCCTGAA AGGAGATACT TCTACTGGTA CACCATCATT GGCTGGGGGA CCCCAACTGT   1080
GTGTGTGACA GTGTGGGCTA CGCTGAGACT CTACTTTGAT GACACAGGCT GCTGGGATAT   1140
GAATGACAGC ACAGCTCTGT GGTGGGTGAT CAAAGGCCCT GTGGTTGGCT CTATCATGGT   1200
TAACTTTGTG CTTTTTATTG GCATTATCGT CATCCTTGTG CAGAAACTTC AGTCTCCAGA   1260
CATGGGAGGC AATGAGTCCA GCATCTACTT GCGACTGGCC CGGTCCACCC TGCTGCTCAT   1320
CCCACTATTC GGAATCCACT ACACAGTATT TGCCTTCTCC CCAGAGAATG TCAGCAAAAG   1380
GGAAAGACTC GTGTTTGAGC TGGGGCTGGG CTCCTTCCAG GGCTTTGTGG TGGCTGTTCT   1440
CTACTGTTTT CTGAATGGTG AGGTACAAGC GGAGATCAAG CGAAAATGGC GAAGCTGGAA   1500
GGTGAACCGT TACTTCGCTG TGGACTTCAA GCACCGACAC CGGTCTCTGG CCAGCAGTGG   1560
GGTGAATGGG GCCACCCAGC TCTCCATCCT GAGCAAGAGC AGCTCCCAAA TCCGCATGTC   1620
TGGCCTCCCT GCTGACAATC TGGCCACCTG AGCCATGCTC CCCT
```

Seq ID NO: 276 <u>Protein sequence</u>
Protein Accession #: NP_001109.1

```
1          11         21         31         41    ·     51
|          |          |          |          |          |
MAGVVHVSLA AHCGACPWGR GRLRKGRAAC KSAAQRHIGA DLPLLSVGGQ WCWPRSVMAG   60
```

```
      TGCGGGAGAG CCTGGCCAAG TGCTGCAGTT GTTCAAGAAA GAGAGCCTTT GGTGTGCTAA    420
      AGACTCTTGT GCCCATCTTG GAGTGGCTCC CCAAATACCG AGTCAAGGAA TGGCTGCTTA    480
5     GTGACGTCAT TTCGGGAGTT AGTACTGGGC TAGTGGCCAC GCTGCAAGGG ATGGCATATG    540
      CCCTACTAGC TGCAGTTCCT GTCGGGATATG GTCTCTACTC TGCTTTTTTC CCTATCCTGA    600
      CATACTTTAT CTTTGGAACA TCAAGACATA TCTCAGTTGG ACCTTTTCCA GTGGTGAGTT    660
      TAATGGTGGG ATCTGTTGTT CTGAGCATGG CCCCCGACGA ACACTTCTC GTATCCAGCA    720
      GCAATGGAAC TGTATTAAAT ACTACTATGA TAGACACTGC AGCTAGAGAT ACAGCTAGAG    780
      TCCTGATTGC CAGTGCCCTG ACTCTGCTGG TTGGAATTAT ACAGTTGATA TTTGGTGGCT    840
10    TGCAGATTGG ATTCATAGTG AGGTACTTGG CAGATCCTTT GGTTGGTGGC TTCACAACAG    900
      CTGCTGCCTT CCAAGTGCTG GTCTCACAGC TAAAGTATTGT CCTCAATGTT TCAACCAAAA    960
      ACTACAATGG AGTTCTCTCT ATTATCTGGA CGCTGGTTGA GATTTTTCAA AATATTGGTG   1020
      ATACCAATCT TGCTGATTTC ACTGCTGGAT TGCTCACCAT TGTCGTCTGT ATGGCAGTTA   1080
      AGGGAATTAAA TGATCGGTTT AGACACAAAA TCCCAGTCCC TATTCCTATA GAAGTAATTG   1140
15    TGACGATAAT TGCTACTGCC ATTTCATATG GAGCCAACCT GGAAAAAAAT TACAATGCTG   1200
      GCATTGTTAA ATCCATCCCA AGGGGGTTTT TGCCTCCTGA ACTTCCACCT GTGAGCTTGT   1260
      TCTCGGAGAT GCTGGCTGCA TCATTTTCCA TCGCTGTGGT GGCTTATGCT ATTGCAGTGT   1320
      CAGTAGGAAA AGTATATGCC ACCAAGTATG ATTACACCAT CGATGGGAAC CAGGAATTCA   1380
      TTGCCTTTGG GATCAGCAAC ATCTTCTCAG GATTCTTCTC TTGTTTTGTG GCCACCACTG   1440
20    CTCTTTCCCG CACGGCCGTC CAGGAGAGCA CTGGAGGAAA GACACAGGTT GCTGGGCATCA   1500
      TCTCTGCTGC GATTGTGATG ATCGCCATTC TTGCCCTGGG GAAGCTTCTG GAACCCTTGC   1560
      AGAAGTCGGT CTTGGCAGCT GTTGTAATTG CCAACCTGAA AGGGATGTTT ATGCAGCTGT   1620
      GTGACATTCC TCGTCTGTGG AGACAGAATA AGATTGATGC TGTTATCTGG GTGTTTACGT   1680
      GTATAGTGTC CATCATTCTG GGGCTGGATC TCGGTTTACT AGCTGGCCTT ATATTTGGAC   1740
25    TGTTGACTGT GGTCCTGAGA GTTCAGTTTC CTTCTTGGAA TGGCCTTGGA AGCATCCCTA   1800
      GCACAGATAT CTACAAAAGT ACCAAGAATT ACAAAAACAT TGAAGAACCT CAAGGAGTGA   1860
      AGATTCTTAG ATTTTCCAGT CCTATTTTCT ATGGCAATGT CGATGGTTTT AAAAAATGTA   1920
      TCAAGTCCAC AGTTGGATTT GATGCCATTA GAGTATATAA TAAGAGGCTG AAAGCGCTGA   1980
      GGAAAATACA GAAACTAATA AAAAGTGGAC AATTAAGAGC AACAAAGAAT GGCATCATAA   2040
30    GTGATGCTGT TTCAACAAAT AATGCTTTTG AGCCTGATGA GGATATTGAA GATCTGGAGG   2100
      AACTTGATAT CCCAACCAAG GAAATAGAGA TTCAAGTGGA TTGGAACTCT GAGCCTTCCAG   2160
      TCAAAGTGAA CGTTCCCAAA GTGCCAATCC ATAGCCTTGT GCTTGACTGT GGAGCTATAT   2220
      CTTTCCTGGA CGTTGTTGGA GTGAGATCAC TGCGGGTGAT TGTCAAAGAA TTCCAAAGAA   2280
      TTGATGTGAA TGTGTATTTT GCATCACTTC AAGATTATGT GATAGAAAAG CTGGAGCAAT   2340
35    GCGGGGTTCTT TGACGACAAC ATTAGAAAGG ACACATTCTT TTTGACGGTC CATGATGCTA   2400
      TACTCTATCT ACAGAACCAA GTGAAATCTC AAGAGGGTCA AGGTTCCATT TTAGAAACGA   2460
      TCACTCTCAT TCAGGATTGT AAAGATACCC TTGAATTAAT AGAAACAGAG CTGACGGAAG   2520
      AAGAACTTGA TGTCCAGGAT GAGGCTATGC GTACACTTGC ATCCTGAAAG TGGGTTCGGG   2580
      AGGTCTCTAT GAGCAAGGAA TACAAGACAA AACTTCCTCA ATGCATTGAC TATTTCTTCA   2640
40    GACTCAAAAC ACTCATTCTT TTTTCTATTA AGCCATTGAA AGAGAAGCAC TAAGACTGCT   2700
      TCTAGGCTTT ATTTATAAAA TAAACACCTT ATCCCTAACA TGGGCAAAAT GGCTAGAATT   2760
      ATTCAGACGA TTTGGCAGCG TCCAGGGTAA GCTGGTGTTA TAATACGCTG CTGATCTACA   2820
      TCACAGATTT GCTAATAATG TTCACGTGGG CCCTGGCATA TCTCTGTTCA GTTAGAGTGA   2880
      GTGCTGACCC AACAGCCTCT GTGGTCAAGC GAGTCACGAA TGATTAATCA TAAAGAAAAA   2940
45    TCAGTTTTTG ACTGACCTGG ATATCCATGA GCTGCACTGA TCACCATGTA AGGTCACATT   3000
      TAGTAAAAGC TGAAATAAAA TGATTAAATGC ATTTATCAAT AAAAGCCTTT GAAAATACTT   3060
      TGGATAATAA ATTGGAGTTT TAAAAATGCA AATTTGCTTA GTATCTAATA ATGAAGTGTT   3120
      ATTACATATA GCCGGAATTG AGGATCTCTT TGATCCTGGA AATGGTTTAC CTAAAAGCTA   3180
      CAGAACCAGG CCAATATATT TTGAAATATT GATGCAGACA AATGAAATAA TAAAGAGATT   3240
50    TTCATGGTTT ATAAAAATCT TTTTTGATAT GATAATAATC ATGATCACAA CTGAGATCAA   3300
      AAAAATATAT GACAGATTAT TTTGTTAAA AATGCAGTTT TAATTATCTT AGTCTATAGA   3360
      AATGATCATT GCATGGAGGC ATGTATAGGT ATGATCTGTG TAAAATCTGA CATAAAAACA   3420
      GTGCTATTCT GAGTGAAAAT TTTTTTGATG TGCTTACATA ACCATGGTGA TTAAAATGAG   3480
      TTTATATTTT TTCTCAAAAA TTTTAGCAGT GTGTAAAGTA AGTAATCTTT AACTGAACTC   3540
55    TGACCACTTA AAAAAAAATC TAAAAATTGA ACTACCTATA GTAGTCTGTG TTTAAAGTGA   3600
      ATTTTTAAAG ACAAAGCATT CTAAATGAAC TCAATATAAA ACATTCATT TGGAATGTAC   3660
      ATACTGAAAA ATACAGGTTT TTTTGACCAA AAGTTTTAT ATCTTTTCTT TTTATTTATT   3720
      TTTTTCCTAA GTGCCAACAA TTTTCTAGAT ATTATATACA ACACAGGCTT TGATCTTGGG   3780
      GACTTTTCCC ATATATTTCA CACTGGAGTG AATGAAGTTG TACTTCATTT CTAGAGAAAA   3840
60    GTTATACCCA GGTCCCCAAT TGAGAATGTC TTGCTTGATT GAAAACGACA TCATCCCTTG   3900
      GTATACTCCA GGGATTGGTT TCAGGACCCC TGCATTACC AAAATTTGTG CACACTCAAG   3960
      TCCTGCAGTC ACCCCTGCCT AAAGATAGAA TGGCTTCTCT GTTTTTCTTC TGAAATACAA   4020
      CCAGAAACAA TGTGTCTATT TCTGAAAGAA TAGGATTAAT GATCATACAA ATGGGTTAAT   4080
      CCTGAATTCT GGTTGTAAAT CTGGTTACAG CATAACTAGG ATTATAATGC TGCCTCATTT   4240
65    TCACAGACAT ACTTGCTTAT ATTGACAACA AATCATCTCG CTAAAGAGTG AATGTAGGCC   4200
      AGGCGCGGTG GCTCATGCCT GTAATCCCAG CACTTTGGGA GGCCGAGGCG GGTGGATCAC   4260
      GAGGTCAGGA GATCGAGACC ATCCTGGCTA ACATGGTAAA ACCCCGTCTC TACTAAAAAT   4320
      AGAAAAAAAG AAATTAGCCT AGCGTGGTGG CTGGCGGGCG CCTGTAGTCC CAGCTATTTG   4380
      GGAGGCTAAG GCAGGAGAAT GGCGTGAACC CGGGAGGCGG AGCTTGCAGT GAGCCGAGGT   4440
70    CGTGCCACTG CACTCCAGCC TGGGCGACAG AGCAAGACTC CGTCTCAAAA AAAAAAAAA   4500
      AAAAAAAAA AGAGTGAATG TAATAGTCTT GCAGAAAATG AATGAATACC TTTGTTCAAT   4560
      AAAGGAAATA TGCACTGCTC ACTTTTTTGA AGGAAATGCC AAAGTTACGT TTTACAACAA   4620
      GGCTAGAGTT TGTAAATTCT GGGTTCATTT GTGATGACAT AAGTCAGCAA ACTGCCGGGA   4680
      TACTGTCTCT TCTATGTATT TTGTGAATAG TAAGCATAAT TTTAGTTTTG TATTATCAAT   4740
75    GAAAATTTCA CTTGAAATTA AAGCTGCCTT TTGTTATATT TTTAACCTAT AGGAATAAGAT   4800
      TCCAGTATTG TATATGAGTT TTAACAAATT AAAAAATCAA ATCATGTACA TTTGAAAATA   4860
      TTTGCACACA TTTAAAAATA AATGTAAAGT TGTCTTTTAA ACTACTCGGA TGTGTCCTTT   4920
      CTGAACAAAA
```

Transmembrane domains: 81-103, 109-131, 136-158, 185-207, 221-243, 245-267, 270-291, 295-317,
                              347-369, 386-408, 420-442, 448-470, 486-508
Cellular Localization: plasma membrane

```
 1          11         21         31         41         51
 |          |          |          |          |          |
MAAPGGRSEP PQLPEYSCSY MVSRPVYSEL AFQQQHERRL QERKTLRESL AKCCSCSRKR    60
AFGVLKTLVP ILEWLPKYRV KEWLLSDVIS GVSTGLVATL QGMAYALLAA VPVGYGLYSA   120
FFPPILTYFIP GTSRHISVGP PPVVSLMVGS VVLSMAPDEH FLVSSSNGTV LNTTMIDTAA   180
RDTARVLIAS ALTLLVGIIQ LIFGGLQIGF IVRYLADPLV GGFTTAAAFQ VLVSQLKIVL   240
NVSTKNYNGV LSIIYTLVSI FQNIGDTNLA DFTAGLLTIV VCMAVKELND RFRHKIPVPI   300
PIEVIVTIIA TAISYGANLE KNYNAGIVKS IPRGFLPPEL PPVSLFSEML AASFSIAVVA   360
YAIAVSVGKV YATKYDYTID GNQEFIAFGI SNIFSGFFSC FVATTALSRT AVQESTGGKT   420
QVAGIISAAI VMIAILALGK LLEPLQKSVL AAVVIANLKG MFMQLCDIPR LWRQNKIDAV   480
IWVFTCIVSI ILGLDLGLLA GLIFGLLTVV LRVQFPSWNG LGSIPSTDIY KSTKNYKNIE   540
EPQGVKILRF SSPIFYGNVD GFKKCIKSTV GFDAIRVYNK RLKALRKIQK LIKSGQLRAT   600
KNGIISDAVS TNNAFEPDED IEDLEELDIP TKEIEIQVDW NSRLPVKVNV PKVPIHSLVL   660
DCGAISFLDV VGVRSLRVIV KEFQRIDVNV YFASLQDYVI EKLEQCGPFD DNIRKDTFFL   720
TVHDAILYLQ NQVKSQEGQG SILETITLIQ DCKDTLELIE TELTEEELDV QDEAMRTLAS   780
QDEAMRTLAS
```

<u>A213 DNA SEQUENCE:</u>
Gene name: ESTs, Highly similar to calcium-activated potassium channel rSK2 [R.norvegicus]
Unigene number: Hs.98280
Probeset Accession #:   AA418000
Nucleic Acid Accession #:   NM_021614
Coding sequence: 458-2197 (underlined sequences correspond to start and stop codons)

```
 1          11         21         31         41         51
 |          |          |          |          |          |
CGGCGGCAGC AGCCCATGCC TCCGGTGCAA CAGCTGCGCC TCCTCCGGTG CCCCGGCGGC    60
GGGGGCGGGA GATAACCTGT CCCTGCTGCT CCGCACCTCC TCGCCCGGCG GCGCCTTCCG   120
GACCCGCACC TCCTCGCCGC TGTCGGGGCTC GTCCTGCTGC TGCTGCTGCT GCTCGTCGCG   180
CCGGGGCAGC CAGCTCAATG TGAGCGAGCT GACGCCGTCC AGCCATGCCA GTGCGCTCCG   240
GCAGCAGTAC GCGCAGCAGT CCGCGCAGCA GTCGGCGTCC GCCTCCCAGT ACCACCAGTG   300
CCACAGCCTG CAGCCGCCG  CCAGCCCCAC GGGCAGCCTC GGCAGTCTGG GCTCCGCGCC   360
CCCGCTCTCG CACCACCACC ACCACCCGCA CCGGGCGCAC CACCAGCACC ACCAGCCCCA   420
GGCGCGCCGC GAGAGCAACC CCTTCACCGA AATAGCCATG AGCAGCTGCA GGTACAACGG   480
GGGCGTCATG CGGCCGCTCA GCAACTTGAG CGCGTCCGC  CGGAACCTCC ACGAGATGGA   540
CTCAGAGGCG CAGCCCCTGC AGCCCCCCGC GTCTGTCGGA GGAGGTGGCG GCGCGTCCTC   600
CCCGTCTGCA GACGCTGCCG CCGCCGCCGC TGTTCGTCC  TCAGCCCCCG AGATCGTGGT   660
GTCTAAGCCC GAGCACAACA ACTCCAACAA CCTGGCGCTC TATGGAACCG GCGGCGGAGG   720
CAGCACTGGA GGAGGCGGCG GCGGTGGGAGG GAGCGGGCAC GGCAGCAGCA GTGGCACCAA   780
GTCCAGCAAA AAGAAAAACC AGAACATCGG CTACAAGCTG GGCCACCGGC GCGCCCTGTT   840
CGAAAAGCGC AAGCGGCTCA GCGACTACGC GCTCATCTTC GGCATGTTCG GCATCGTGGT   900
CATGGTCATC GAGACCGAGC TGTCGTGGGG CGCCTACGAC AAGGCGTGC  TGTATTCCTT   960
AGCTCTGAAA TGCCTTATCA GTCTCTCCAC GATCATCCTG CTCGGTCTGA TCATCGTGTA  1020
CCACGCCAGG GAAATACAGT TGTTCATGGT GGACAATGGA GCAGATGACT GGAGAATAGC  1080
CATGACTTAT GAGCGTATTT TCTTCATCTG CTTGGAAATA CTGGTGTGTG CTATTCATCC  1140
CATACCTGGG AATTATACAT TCACATGGAC GGCCCGGCTT GCCTTCTCCT ATGCCCCATC  1200
CACAACCACC GCTGATGTGG ATATTATTTT ATCTATACCA ATGTTCTTAA GACTCTATCT  1260
GATTGCCAGA GTCATGCTTT TACATAGCAA ACTTTTCACT GATGCCTCCT CTAGAAGCAT  1320
TGGAGCACTT AATAAGATAA ACTTCAATAC ACGTTTTGTT ATGAAGACTT TAATGACTAT  1380
ATGCCCAGGA ACTGTACTCT TGGTTTTTAG TATCTCATTA TGGATAATTG CCGCATGGAC  1440
TGTCCGAGCT TGTGAAAGGT ACCATGATCA ACAGGATGTT ACTAGCAACT TCCTTGGAGC  1500
GATGTGGTTG ATATCAATAA CTTTTCTCTC CATTGGTTAT GGTGACATGG TACCTAACAC  1560
ATACTGTGGA AAAGGAGTCT GCTTACTTAC TGGAATTATG GGTGCTGGTT GCACAGCCCT  1620
GGTGGTAGCT GTAGTGGCAA GGAAGCTAGA ACTTACCAAA GCAGAAAAAC ACGTGCACAA  1680
TTTCATGATG GATACTCAGC TGACTAAAAG AGTAAAAAAT GCCGCTGCCA ATGTACTCAG  1740
GGAAACATGG CTAATTTACA AAAATACAAA GCTAGTGAAA AAGATAGATC ATGCAAAAGT  1800
AAGAAAACAT CAACGAAAAT TCCTGCAAGC TATTCATCAA TTAAGAAGTG TAAAAATGGA  1860
GCAGAGGAAA CTGAATGACC AAGCAAACAC TTTGGTGGAC TTGGCAAAGA CCCAGAACAT  1920
CATGTATGAT ATGATTTCTG ACTTAAACGA AAGGAGTGAA GACTTCGAGA AGAGGATTGA  1980
TACCCTGGAA ACAAAACTAG AGACTTTGAT TGGTAGCATC CACGCCCTCC CTGGGCTCAT  2040
AAGCCAGACC ATCAGGCAGC AGCAGAGAGA TTTCATTCAG GCTCAGATGG AGAGCTACGA  2100
CAAGCACGTC ACTTACAATG CTGAGCGGTC CCGGTCCTCG TCCAGGAGGC GGCGGTCCTC  2160
TTCCACAGCA CCACCAACTT CATCAGAGAG TAGCTAGAAG AGAATAAGTT AACCACAAAA  2220
TAAGACTTTT TGCCATCATA TGGTCAATAT TTTAGCTTTT ATTGTAAAGC CCCTATGGTT  2280
CTAATCAGCG TTATCCGGGT TCTGATGTCA GAATCCTGGG AACCTGAACA CTAAGTTTTA  2340
GGCCAAAATG AGTGAAAACT CTTTTTTTTT CTTTCAGATG CACAGGGAAT GCACCTATTA  2400
TTGCTATATA GATTGTTCCT CCTGTAATTT CACTAACTTT TTATTCATGC ACTTCAAACA  2460
AACTTTACTA CTACATTATA TGATATATAA TAAAAAAAGT TAATTTCGGA
```

<u>A214 Protein sequence:</u>
Gene name: ESTs, Highly similar to calcium-activated potassium channel rSK2 [R.norvegicus
Unigene number: Hs.98280
Probeset Accession #:   AA418000
Protein Accession #:   NP_067627
Signal sequence: none found
Transmembrane domains: 135-157, 168-190, 208-230, 254-276, 306-328, 342-364, 373-394

Calmodulin binding domain: 412-488
Cellular Localization: plasma membrane

```
          1         11        21        31        41        51
          |         |         |         |         |         |
        MSSCRYNGGV MRPLSNLSAS RRNLHEMDSE AQPLQPPASV GGGGGASSPS ADAAAAAAVS    60
        SSAPEIVVSK PEHNNSNNLA LYGTGGGGST GGGGGGGGSG HGSSSGTKSS KKKNQNIGYK   120
        LGHRRALFEK RRRLSDYALI PGMFGIVVMV IETELSWGAY DKASLYSLAL KCLISLSTII   180
        LLGLIIVYHA REIQLFMVDN GADDWRIAMT YERIFFICLE ILVCAIHPIP GNYTFTWTAR   240
        LAFSYAPSTT TADVDIILSI PMFLRLYLIA RVMLLHSKLF TDASSRSIGA LNKINFNTRP   300
        VMKTLMTICP GTVLLVFSIS LWIIAAWTVR ACSRYHDQQD VTSNFLGAMW LISITPLSIG   360
        YGDMVPNTYC GKGVCLLTGI MGAGCTALVV AVVARKLELT KAEKHVHNFM MDTQLTKRVK   420
        NAAANVLRET WLIYKNTKLV KKIDHAKVRK HQRKFLQAIH QLRSVKMEQR KLNDQANTLV   480
        DLAKTQNIMY DMISDLNERS EDFEKRIVTL ETKLETLIGS IHALPGLISQ TIRQQQRDFI   540
        EAQMESYDKH VTYNAERSRS SSRRRRSSST APPTSSESS
```

A215 DNA SEQUENCE:

Gene name:	CGI-86 protein
Unigene number #:	Hs.109201
Probeset Accession #:	AW161450
Nucleic Acid Accession #:	NM_016029
Coding sequence:	228-1097 (underlined sequences correspond to start and stop codons)

```
          1         11        21        31        41        51
          |         |         |         |         |         |
        CTGCGATCCC GCAGGGCAGC GACGCGACTC TGGTGCGGGC CGTCTTCTTC CCCCCGAGCT     60
        GGGCGTGCGC GGCCGCAATG AACTGGGAGC TGCTGCTGTG GCTGCTGGTG CTGTGCGCGC    120
        TGCTCCTGCT CTTGGTGCAG CTGCTGCGCT TCCTGAGGGC TGACGGCGAC CTGACGCTAC    180
        TATGGGCCGA GTGGCAGGGA CGACGCCCAG AATGGGAGCT GACTGATATG GTGGTGTGGG    240
        TGACTGGAGC CTCGAGTGGA ATTGGTGAGG AGCTGGCTTA CCAGTTGTCT AAACTAGGAG    300
        TTTCTCTTGT GCTGTCAGCC AGAAGAGTGC ATGAGCTGGA AAGGGTGAAA AGAAGATGCC    360
        TAGAGAATGG CAATTTAAAA GAAAAAGATA TACTTGTTTT GCCCCTTGAC CTGACCGACA    420
        CTGGTTCCCA TGAAGCGGCT ACCAAAGCTG TTCTCCAGGA GTTTGGTAGA ATCGACATTC    480
        TGGTCAACAA TGGTGGAATG TCCCAGCCGTT CTCTGTGCAT CTTGGTGTCT              540
        ACAGAAAGCT AATAGAGCTT AACTACTTAG GGACGGTGTC CTTGACAAAA TGTGTTCTGC    600
        CTCACATGAT CGAGAGGAAG CAAGGAAAGA TTGTTACTGT GAATAGCATC CTGGGTATCA    660
        TATCTGTACC TCTTTCCATT GGATACTGTG CTAGCAAGCA TGCTCTCCGG GGTTTTTTTA    720
        ATGGCCTTCG AACAGAACTT GCCACATACC CAGGTATAAT AGTTTCTAAC ATTTGCCCAG    780
        GACCTGTGCA ATCAAATATT GTGGAGAATT CCCTAGCTGG AGAAGTCACA AAGACTATAG    840
        GCAATAATGG AGACCAGTCC CACAAGATGA CAACCAGTCG TTGTGTGCGG CTGATGTTAA    900
        TCAGCATGGC CAATGATTTG AAAGAAGTTT GGATCTCAGA ACAACCTTTC TTGTTAGTAA    960
        CATATTTGTG GCAATACATG CCAACCTGGG CCTGGTGGAT AACCAACAAG ATGGGGAAGA   1020
        AAAGGATTGA GAACTTTAAG AGTGGTGTGG ATGCAGACTC TTCTTATTTT AAAATCTTTA   1080
        AGACAAAACA TGACTGAAAA GAGCCACCTGT ACTTTTCAAG CCACTGGAGG GAGAAATGGA   1140
        AAACATGAAA ACAGCCAATCT TCTTATGCTT CTGAATAATC AAAGACTAAT TTGTGATTTT   1200
        ACTTTTTAAT AGATATGACT TTGCTTCCAA CATGGAATGA AATAAAAAAT AAATAATAAA   1260
        AGATTGCCAT GAATCTTGCA AA
```

A216 Protein sequence:

Gene name:	CGI-86 protein
Unigene number:	Hs.109201
Probeset Accession #:	AW161450
Protein Accession #:	NP_057113
Signal sequence:	1-26
Transmembrane domains:	183-206, 221-243
Cellular Localization:	plasma membrane

```
          1         11        21        31        41        51
          |         |         |         |         |         |
        MNWELLLWLL VLCALLLLLV QLLRFLRADG DLTLLWAEHQ GRRPEWELTD MVVWVTGASS     60
        GIGEELAYQL SKLGVSLVLS ARRVHELERV KRRCLENGNL KEKDILVLPL DLTDTGSHEA    120
        ATKAVLQEFG RIDILVNNGG MSQRSLCMDT SLDVYRKLIE LNYLGTVSLT KCVLPHMIER    180
        KQGKIVTVNS ILGIISVPLS IGYCASKHAL RGFFNGLRTE LATYPGIIVS NICPGPVQSN    240
        IVENSLAGEV TKTIGNNGDQ SHKMTTSRCV RLMLISMAND LKEVWISEQP FLLVTYLWQY    300
        MPTWAWWITN KMGKKRIENF KSGVDADSSY FKIFKTKHD
```

A217 DNA SEQUENCE:

Gene name:	Homo sapiens mRNA; cDNA DKFZp56401763 (from clone DKFZp56401763)
Unigene number:	Hs.27373
Probeset Accession #:	F13036
Nucleic Acid Accession #:	AC012478
Coding sequence:	1-1896 (underlined sequences correspond to start and stop codons)

```
          1         11        21        31        41        51
          |         |         |         |         |         |
        ATGCGCGCCG TGCCGCTGCC CGCCCCGCTC CTGCCGCTGC TGCTGCTCGC GCTCCTGGCC     60
        GCTCCCGCCG CCCGCGCCAG CAGAGCCGAG TCCGTCTCCG CGCCGTGGCC CGAACCCGAG    120
        CGCGAGTCGC GGCCACCGCC CGGCCCGGGG CCCGGGAACA CCACCCGGTT TGGGTCTGGG    180
        GCGGCGGGCG GCAGCGGCAG CTCCAGCTCC AACAGCAGTG GCGACGCCTT GGTGACCCGC    240
```

```
ATTTCCATCC TCCTCCGCGA CCTACCCACC CTCAAGGCAG CCGTGATCGT GGCGTTCGCC    300
TTTACCACCC TCCTCATCGC CTGCCTGCTG CTGCGCGTCT TCAGGTCGGG AAAGAGGTTA    360
AAGAAGACAC GCAAGTATGA TATCATCACC ACTCCAGCAG AGCGAGTGGA AATGGCGCCA    420
CTAAATGAAG AGGATGATGA AGATGAGGAC TCCACAGTAT TCGACATCAA ATACAGAGTG    480
TCCTTGCCGG CTGCACTGAG ACGTCAGCTG CCAGGGTGCC AGACGCTACT GACAGTTCCT    540
GTGCCCCCAC CCTTCATCCT CGACATTGAC CTTCCAGCAA GATGCAGTGG AAGGCCTGAT    600
GGTGGAATCA GACCTGGTAA AACCTGTTTC CCAGCCTGGT GGCATCCTGT GGAAAGTTGG    660
TCAGCTGCAA CCTGGGGTGT GAAGGACTGG ACCTGGAAGC CCTCTTGCGT CGGAGGTGTT    720
GAAACCAAAA CGAACGTTAT GTATAAAACC CCAGCTCCAT CGTGCGTGTC AGGCATCTGC    780
TCAGACTGTC ACTGGCAAGC TCGTTTCCAC GTCACCACAA TGGAGTTGCT TCTGCCACCC    840
TTTGGGCATC CCTTTAAAGT GCCCCCTACT TCTACTCCCC ATGGTTTTCG ACAACTGCAG    900
CTGAATCTCA TGGAAAAGCT GGATTCCTCT GCCTTACGCA GAAACACCCG GGCTCCATCT    960
GCCAGGTGCT TGCCACTGGT CCTGGCAGAA ATGGCGGCTG CTGAAAGTGA CCTTCCAAAT   1020
CCTTGGTGGC ACTTCAGCGC CACAGGCTCT CCAATAAAAA CCCTTTACAC ACAAACCATG   1080
AGTACCTTGG GCTTGGATGT TTTCTGTGGT GCCGGCCAGC GGGGCACCTT TTGTGAAGAC   1140
AGAGCAGTGA CTAAGGTTCT CCAGGGTAGC TCTTTCTCCA AACAGCTGCG CTGGAAGCCA   1200
GCCCTAGAGA GTGGGTTTCC CCATCATCTC AGGCTTCTCA GAGAGTGTCC TCCGCTGAGC   1260
ACCCATCCTG TCAGGTTGGC TCGTTCAGAT GCCCGGGGAC AAGCCAGCCT GACGGGGAGG   1320
AGGGTGTTTC GGCGTCCGCG GCAGTCTCTG CATGGCGGAG GGTCAGCGGG TACCGCAACT   1380
TGCCTTTTGG TTTTGAAGAT TCTGTTGAGG CGCCATCCTC ACCTTGACCT CTTCTACAAA   1440
ATCTGTCTCC CCTGCTGTGC CGTGGAACAC CTACGGGAAG CCAAGAGAAG CTCAGTGACT   1500
GTCCTTGCGT CATTTGAGCA GAGCCCACAA AAGGCAGCTG CTGCCCACGG GGAGCCTGTC   1560
AAACGAGGGC CCAGTGGGCA ATTGACCAGA CACACATGCC CTGGCTGGGG GATCACACAT   1620
GCGAACCTGC AGACAATTCC AGATACCCAA GGCCAGGAAG GCCCACGTGA GGATGTCACT   1680
CACCCTGGAG GAGACTTGGA TGGGGTGGCA AATTTCTATT TGGAGGAAGA GGGTTTCCAG   1740
GATGGCAGAT GCCAGAAGAT GGTCCTGATG TCTGAGGAAG GGCCACCTAG TTTGACAGGA   1800
TGTGAGAGGC TCACAGGTTC CCATCACTTC TCCAGCCATT CCAAGTCTTG GTCCTTCCTT   1860
TCCCCCCGAC AGCCCCTGTT TCTGTCCAGG CCCTGA
```

A218 Protein sequence:

Gene name:	Homo sapiens mRNA; cDNA DKFZp56401763 (from clone DKFZp56401763)
Unigene number:	Hs.27373
Probeset Accession #:	F13036
Protein Accession #:	FGENESH predicted
Signal sequence:	1-27
Transmembrane domains:	94-115, 448-469
Cellular Localization:	not determined

```
1          11         21         31         41         51
|          |          |          |          |          |
MRAVPLPAPL LPLLLLALLA APAARASRAE SVSAFWPEPE RESRPPPGFG PGNTTRFGSG    60
AAGGSGSSSS NSSGDALVTR ISILLRDLPT LKAAVIVAFA FTTLLIACLL LRVFRSGKRL   120
KKTRKYDIIT TPAERVEMAP LNEEDDEDED STVFDIKYRV SLPAALRRQL PGCQTLLTVP   180
VPPPFILDID LPARCSGRPD GGIRPGKTCF PAWWHPVESW SAATWGVKDW TWKPSCVGGV   240
ETKTNVMYKT PAPSCVSGIC SDCHWQARFE VTTMELLLPP FGHPFKVPPT STPHGFRQLQ   300
LNLMEKLDSS ALRRNTRAPS ARCLFLVLAE MAAAESDLPN PWNHFEATGS PIKTLYTQTM   360
STLGLDVFCG AGQRGTFCED RAVTKVLQGS SPSKQLRWKP ALESGFPHHL RLLRECPPLS   420
THPVRLARSD ARGQASLTGR RVFRRPRQSL HGGGSAGTAT CLLVLKILLR RHPELDLFYK   480
ICLPCCAVEH LREAKRSSVT VLASFEQSPQ KAAAAHGEPV KRGPSGQLTR HTCFGWGITH   540
ANLQTIPDTQ GQEGPREDVT HPGGDLDGVA NFYLEEEGFQ DGRCQKMVLM SEEGPPSLTG   600
CERLTGSHHF SSESKSWSFL SPRQPLFLSR P
```

A219 DNA SEQUENCE

Gene name:	selectin E (endothelial adhesion molecule 1)
Unigene number:	Hs.89546
Probeset Accession #:	M24736
Nucleic Acid Accession #:	NM_000450
Coding sequence:	1-1833 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGATTGCTT CACAGTTTCT CTCAGCTCTC ACTTTGGTGC TTCTCATTAA AGAGAGTGGA    60
GCCTGGTCTT ACAACACCTC CACGGAAGCT ATGACTTATG ATGAGGCCAG TGCTTATTGT    120
CAGCAAAGGT ACACACACCT GGTTGCAATT CAAAACAAAG AAGAGATTGA GTACCTAAAC    180
TCCATATTGA GCTATTCACC AAGTTATTAC TGGATTGGAA TCAGAAAAGT CAACAATGTG    240
TGGGTCTGGG TAGGAACCCA GAAACCTCTG ACAGAAGAAG CCAAGAACTG GGCTCCAGGT    300
GAACCCAACA ATAGGCAAAA AGATGAGGAC TGCGTGGAGA TCTACATCAA GAGAGAAAAA    360
GATGTGGGCA TGTGGAATGA TGAGAGGTGC AGCAAGAAGA AGCTTGCCCT ATGCTACACA    420
GCTGCCTGTA CCAATACATC CTGCAGTGGC CACGGTGAAA CCATCAATAA TGGAAGAATAAT    480
TACACTTGCA AGTGTGACCT TGGCTTCAGT GGACTCAAGT GTGAGCAAAT TGTGAACTGT    540
ACAGCCCTGG AATCCCCTGA GCATGGAAGC CTGGTTTGCA GTCACCCACT GGGAAACTTC    600
AGCTACAATT CTTCCTGCTC TATCAGCTGT GATAGGGGTT ACCTGCCAAG CAGCATGGAG    660
ACCATGCAGT GTATGTCCTC TGGAGAATGG AGTGCTCCTA TTCCAGCCTG CAATGTGGTT    720
GAGTGTGATG CTGTGACAAA TCCAGCCAAT GGGTTCGTGG AATGTTCCTA AAACCCTGGA    780
AGCTTCCCAT GGAACACAAC CTGTACATTT GACTGTGAAG AAGGATTTGA ACTAATGGGA    840
GCCCAGAGCC TTCAGTGTAC CTCATCTGGG AATTGGGACA ACGAGAAGCC AACGTGTAAA    900
GCTGTGACAT GCAGGGCCGT CCGCCAGCCT CAGAATGGCT CTGTGAGGTG CAGCCATTCC    960
CCTGCTGGAG AGTTCACCTT CAAATCATCC TGCAACTTCA CCTGTGAGGA AGGCTTCATG   1020
TTGCAGGGAC CAGCCCAGGT TGAATGCACC ACTCAAGGGC AGTGGACACA GCAAATCCCA   1080
GTTTGTGAAG CTTTCCAGTG CACAGCCTTG TCCAACCCCG AGCGAGGCTA CATGAATTGT   1140
CTTCCTAGTG CTTCTGGCAG TTTTCGTTAT GGGTCCAGCT GTGAGTTCTC CTGTGAGCAG   1200
GGTTTTGTGT TGAAGGGATC CAAAAGGCTC CAATGTGGCC CCACAGGGGA GTGGGACAAC   1260
```

```
GAGAAGCCCA CATGTGAAGC TGTGAGATGC GATGCTGTCC ACCAGCCCCC GAAGGGTTTG   1320
GTGAGGTGTG CTCATTCCCC TATTGGAGAA TTCACCTACA AGTCCTCTTG TGCCTTCAGC   1380
TGTGAGGAGG GATTTGAATT ATATGGATCA ACTCAACTTG AGTGCACATC TCAGGGACAA   1440
TGGACAGAAG AGGTTCCTTC CTGCCAAGTG GTAAAATGTT CAAGCCTGGC AGTTCCGGGA   1500
AAGATCAACA TGAGCTGCAG TGGGGAGCCC GTGTTTGGCA CTGTGTGCAA GTTCGCCTGT   1560
CCTGAAGGAT GGACGCTCAA TGGCTCTGCA GCTCGGACAT GTGGGAGCCAC AGGACACTGG   1620
TCTGGCCTGC TACCTACCTG TGAAGCTCCC ACTGAGTCCA ACATTCCCTT GGTAGCTGGA   1680
CTTTCTGCTG CTGGACTCTC CCTCCTGACA TTAGCACCAT TTCTCCTCTG GCTTCGGAAA   1740
TGCTTACGGA AAGCAAAGAA ATTTGTTCCT GCCAGCAGCT GCCAAAGCCT TGAATCAGAC   1800
GGAAGCTACC AAAAGCCTTC TTACATCCTT TAA                                1800
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MIASQFLSAL TLVLLIKESG AWSYNTSTEA MTYDEASAYC QQRYTHLVAI QNKEEIEYLN   60
SILSYSPSYY WIGIRKVNNV WVWVGTQKPL TEEAKNWAPG EPNNRQKDED CVEIYIKREK   120
DVGMWNDERC SKKKLALCYT AACTNTSCSG HGECVETINN YTCKCDPGFS GLKCEQIVNC   180
TALESPEHGS LVCSHPLGNF SYNSSCSISC DRGYLPSSME TKQCMSSGEW SAPIPACNVV   240
ECDAVTNPAN GFVECFQNPG SFPWNTTCTF DCEBGFELMG AQSLQCTSSG NWDNEKPTCK   300
AVTCRAVRQP QNGSVRCSHS PAGEFTFKSS CNFTCEEGFM LQGPAQVECT TQGQWTQQIP   360
VCEAFQCTAL SNPERGYMNC LPSASGSFRY GSSCEFSCEQ GFVLKGSKRL QCGPTGEWDN   420
EKPTCEAVRC DAVHQPPKGL VRCAHSPIGE FTYKSSCAFS CEEGFELYGS TQLECTSQGQ   480
WTEEVPSCQV VKCSSLAVPG KINMSCSGEP VFGIVCKFAC PEGWTLNGSA ARTCGATGHW   540
SGLLPTCEAP TESNIPLVAG LSAAGLSLLT LAPFLLWLRK CLRKAKKFVP ASSCQSLESD   600
GSYQKPSYIL
```

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGATGTGTG AAGTGATGCC CACGATTAAT GAGGACACCC CAATGAGCCA AAGGGGGTCC   60
CAAAGCAGTG GCTCGGACTC AGACTCCCAT TTTGAGCAGC TGATGGTGAA TATGCTAGAT   120
GAAAGGGATC GTCTTCTAGA CACCCTTCGG GAGACCCAGG AAAGCCTCTC ACTTGCCCAG   180
CAAAGACTTC AGGATGTCAT CTATGACCGA GACTCACTCC AGAGACAGCT CAATTCAGCC   240
CTGCCAGACG ATATCGAATC CCTAACAGGA GGGCTGGCTG GTTCTAAGGG GGCTGATCCA   300
CCGGAATTTG CTGCACTGAC AAAAGAATTA AATGCCTGCA GGGAACAACT TCTAGAAAAG   360
GAAGAAGAAA TCTCTGAACT TAAAGCTGAA AGAAACAACA CAAGACTATT ACTGGAGCAT   420
TTGGAGTGCC TTGTGTCACG ACATGAAAGA TCACTAAGAA TGACGGTGGT AAAACGGCAA   480
GCCCAGTCTC CCTCAGGAGT ATCCAGTGAA GTTGAAGTTC TCAAGGCACT GAAATCTTTG   540
TTTGAGCACC ACAAGGCCTT GGATGAAAAG GTAAGGGAGC GACTGAGGGT TTCTTTAGAA   600
AGAGTCTCTG CACTGGAAGA AGAACTAGCT GCTGCTAATC AGGAGATTGT TGCCTTGCGT   660
GAACAAAATG TTCATATACA AAGAAAAATG GCATCAAGCG AGGGATCCAC AGAGTCAGAA   720
CATCTTGAAG GGATGGAACC TGGACAGAAA GTCCATGAGA AGCGTTTGTC CAATGGTTCT   780
ATAGACTCAA CCGATGAAAC TAGTCAAATA GTTGAACTAC AAGAATTGCT TGAAAAGCAA   840
AACTATGAAA TGGCCCAGAT GAAAGAACGT TTAGCAGCCC TTTCTTCCCG AGTGGGAGAG   900
GTGGAACAGG AAGCAGAGAC AGCAAGAAAG GATCTCATTA AAACAGAAGA AATGAACACC   960
AAGTATCAAA GGGACATTAG GGAGGCCATG GCACAAAAGG AAGATATGGA AGAAAGAATT   1020
ACAACCCTTG AAAAGCGTTA CCTCAGTGCT CAGAGAGAAT CTACCTCCAT ACATGACATG   1080
AATGATAAAC TAGAAAATGA GTTAGCAAAT AAAGAAGCTA TCCTACGGCA GATGGAAGAG   1140
AAAAACAGAC AGTTACAAGA ACGTCTTGAG CTAGCTGAAC AAAAGTTGCA GCAGACCATG   1200
AGAAAGGCTG AAACCTTGCC TGAAGTAGAG GCTGAACTGG CTCAGAGAAT TGCAGCCCTA   1260
ACCAAGGCTG AAGAGAGACA TGGAAATATT GAAGAACGTA TGAGACATTT AGAGGGTCAA   1320
CTTGAAGAGA AGAATCAAGA ACTTCAAAGA GCTAGGCAAA GAGAGAAAT GAATGAGGAG   1380
CATAACAAGA GATTATCGGA TACGGTTGAT AGACTTCTGA CTGAATCCAA TGAACGCCTA   1440
CAACTACACT TAAAGGAAAG AATGGCTGCT CTAGAAGAAA AGAATGTTTT AATTCAAGAA   1500
TCAGAAACTT TCAGAAAGAA TCTTGAAGAA TCTTTACATG ATAAGGAAAG ATTAGCAGAA   1560
GAAATTGAAA AGCTGAGATC TGAACTTGAC CAATTGAAAA TGAGAACTGG CTCTTTAATT   1620
GAACCCACAA TACCAAGAAC TCATCTAGAC ACCTCAGCTG AGTTGCGGTA CTCAGTGGGA   1680
TCCCTAGTGG ACAGCCAGTC TGATTACAGA ACAACTAAAG TAATAAGAAG ACCAAGGAGA   1740
GGCCGCATGG GTGTGCGAAG AGATGAGCCA AAGGTGAAAT CTCTTGGGGA TCACGAGTGG   1800
AATAGAACTC AACAGATTGG AGTACTAAGC AGCCACCCTT TTGAAAGTGA CACTGAAATG   1860
TCTGATATTG ATGATGATGA CAGAGAAACA ATTTTTAGCT CAATGGATCT TCTCTCTCCA   1920
AGTGGGTCATT CCGATGCCCA GACGCTAGCC ATGATGCTTC AGGAACAATT GGATGCCATC   1980
AACAAAGAAA TCAGGCTAAT TCAGGAAGAA AAAGAATCTA CAGAGTTGCG TGCTGAAGAA   2040
ATTGAAAATA GAGTGGCTAG TGTGAGCCTC GAAGGCCTGA ATTTGGCAAG GGTCCACCCA   2100
```

```
GGTACCTCCA TTACTGCCTC TGTTACAGCT TCATCGCTGG CCAGTTCATC TCCCCCCAGT  2160
GGACACTCAA CTCCAAAGCT CACCCCTCGA AGCCCTGCCA GGGAAATGGA TCGGATGGGA  2220
GTCATGACAC TGCCAAGTGA TCTGAGGAAA CATCGGAGAA AGATTGCAGT TGTGGAAGAA  2280
GATGGTCGAG AGGACAAAGC AACAATTAAA TGTGAAACTT CTCCTCCTCC TACCCCTAGA  2340
GCCCTCAGAA TGACTCACAC TCTCCCTTCT TCCTACCACA ATGATGCTCG AAGTAGTTTA  2400
TCTGTCTCTC TTGAGCCAGA AAGCCTCGGG CTTGGTAGTG CCAACAGCAG CCAAGACTCT  2460
CTTCACAAAG CCCCCAAGAA GAAAGGAATC AAGTCTTCAA TAGGACGTTT GTTTGGTAAA  2520
AAAGAAAAAG CTCGACTTGG GCAGCTCCGA GGCTTTATGG AGACTGAAGC TGCAGCTCAG  2580
GAGTCCCTGG GGTTAGGCAA ACTCGGAACT CAAGCTGAGA AGGATCGAAG ACTAAAGAAA  2640
AAGCATGAAC TTCTTGAAGA AGCTCGGAGA AAGGGATTAC CTTTTGCCCA GTGGGATGGG  2700
CCAACTGTGG TCGCATGGCT AGAGCTTTGG TTGGGAATGC CTGCGTGGTA CGTGGCAGCC  2760
TGCCGAGCCA ACGTGAAGAG TGGTGCCATC ATGTCTGCTT TATCTGACAC TGAGATCCAG  2820
AGAGAAATTG GAATCAGCAA TCCACTGCAT CGCTTAAAAC TTCGATTAGC AATCCAGGAG  2880
ATGGTTTCCC TAACAAGTCC TTCAGCTCCT CCAACATCTC GAACTCCTTC AGGCCAACGTT  2940
TGGGTGACTC ATGAAGAAAT GGAAAATCTT GCAGCTCCAG CAAAAACGAA AGAATCTGAG  3000
GAAGGAAGCT GGGCCCAGTG TCCGGTTTTT CTACAGACCC TGGCTTATGG AGATATGAAT  3060
CATGAGTGGA TTGGAAATGA ATGGCTTCCC AGCTTGGGGT TACCTCAGTA CAGAAGTTAC  3120
TTTATGGAAT GCTTGGTAGA TGCAAGAATG TTAGATCACC TAACAAAAAA AGATCTCCGT  3180
GTCCATTTAA AAATGGTGGA TAGTTTCCAT CGAACAAGTT TACAATATGG AATTATGTGC  3240
TTAAAGAGGT TGAATTATGA CAGAAAAGAA CTAGAAAGAA GACGGGAAGC AAGCCAACAT  3300
GAAATAAAAG ACGTGTTGGT GTGGAGCAAT GACCGAATTA TTCGCTGGAT ACAAGCAATT  3360
GGACTTCGAG AATATGCAAA TAATATACTT GAGAGCGGTG TGCATGGCTC ACTTATAGCC  3420
CTGGATGAAA ACTTTGACTA CAGCAGCTTA ACTTATTAT TACAGATTCC AACACAGAAC  3480
ACCCAGGCAA GGCAGATTCT TGAAAGAGAA TACAATAACC TCTTGGCCCT GGGAACTGAA  3540
AGGCGACTGG ATGAAAGTGA TGACAAGAAC TTCAGACGTG GATCAACCTG GAGAAGGCAG  3600
TTTCCTCCTC GTGAAGTACA TGGAATCAGC ATGATGCCTG GGTCCTCAGA AACATTACCA  3660
GCTGGATTTA GGTTAACCAC AACCTCTGGG CAATGTCAAGAA AAATGACAAC AGATGTTGCT  3720
TCATCAAGAC TGCAGAGGTT AGACAACTCC ACTGTTCGCA CATACTCATG TCTCGAGTAA  3780
GCGGCCGCTT TAA
```

A222 Protein sequence:

Gene name:	ESTs; Liprin A2
Unigene number:	Hs.306480
Probeset Accession #:	N51002
Protein Accession #:	none found
Signal sequence:	none found
Transmembrane domains:	none found
AAA domain:	286-539
SAM domains:	895-964, 1017-1084, 1105-1177
Cellular Localization:	not determined

```
1          11         21         31         41         51
MMCEVMPTIN EDTPMSQRGS QSSGSDSDSH PEQLMVNMLD ERDRLLDTLR ETQESLSLAQ   60
QRLQDVIYQR DSLQRQLNSA LPQDIESLTG GLAGSKGADP PEFAALTKEL NACREQLLEK  120
EEEISELKAE RNNTRLLLEH LECLVSRHER SLRMTVVKRQ AQSPSGVSSE VEVLKALKSL  180
FEHHKALDEK VRERLRVSLE RVSALEEELA AANQEIVALR EQNVHIQRKM ASSEGSTESE  240
HLEGMEPGQK VHEKRLSNGS IDSTDETSQI VELQELLEKQ NYEMAQMKER LAALSSRVGE  300
VEQEAETARK DLIKTEEMNT KYQRDIREAM AQKEDMEERI TTLEKRYLSA QRESTSIHDM  360
NDKLENELAN KEAILRQMEE KNRQLQERLE LAEQKLQQTM RKAETLPEVE AELAQRIAAL  420
TKAEERHGNI EERMRHLEGQ LEEKNQELQR ARQRERNNEE HNKRLSDTVD RLLTESNERL  480
QLHLKERMAA LEEKNVLIQE SETFRKNLEE SLHDKERLAE EIEKLRSELD QLKMRTGSLI  540
EPTIPRTHLD TSAELRYSVG SLVDSQSDYR TTKVIRRPRR GRMGVRRDEP KVKSLGDHEW  600
NRTQQIGVLS SEPFESDTEM SDIDDDDRET IFSSMDLLSP SGHSDAQTLA MMLQEQLDAI  660
NKEIRLIQEE KESTELRAEE IENRVASVSL EGLMLARVEP GTSITASVTA SSLASSSPPS  720
GHSTPKLTPR SPAREMDRMG VMTLPSDLRK HRRKIAVVEE DGREDKATIK CSTSPPPTPR  780
ALRMTHTLPS SYHNDARSSL SVSLEPESLG LGSANSSQDS LHKAPKKKGI KSSIGRLFGK  840
KEKARLGQLR GFMSTEAAAQ ESLGLGKLGT QAEKDRRLRK KHELLEEARR KGLPFAQWDG  900
PTVVAWLELW LGMPAWYVAA CRANVKSGAI MSALSDTEIQ REIGISNPLH RLKLRLAIQE  960
MVSLTSPSAP PTSRTPSGNV WVTHEEMENL AAPAKTKESE EGSWAQCPVP LQTLAYGDMN 1020
HEWIGNEWLP SLGLPQYRSY FMECLVDARM LDHLTKKDLR VRLKMVDSFH RTSLQYGIMC 1080
LKRLNYDRKE LERRREASQH EIKDVLVWSN DRIIRWIQAI GLREYANNIL EEGVHGSLIA 1140
LDENFDYSSL TLLLQIPTQN TQARQILERE YNNLLALGTE RRLDESDDKN FRRGSTWRRQ 1200
FPPREVHGIS MMPGSSETLP AGFRLTTTSG QSRKMFTDVA SSRLQRLDNS TVRTYSCLE
```

A223 DNA SEQUENCE

Gene name:	CDA14
Unigene number:	Hs.26813
Probeset Accession #:	N32912
Nucleic Acid Accession #:	NM_016570
Coding sequence:	1- 1134 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
ATGAGGCGAC TGAATCGGAA AAAAACTTTA AGTTTGGTAA AAGAGTTGGA TGCCTTTCCG   60
AAGGTTCCTG AGAGCTATGT AGAGACTTCA GCCAGTGGAG GTACAGTTTC TCTAATAGCA  120
TTTACAACTA TGGCTTTATT AACCATAATG GAATTCTCAG TATATCAAGA TACATGGATG  180
AAGTATGAAT ACGAAGTAGA CAAGGATTTT TCTAGCAAAT TAAGAATTAA TATAGATATT  240
ACTGTTGCCA TGAAGTGTCA ATATGTTGGA GCGGATGTAT TGGATTTAGC AGAAACAATG  300
GTTGCATCTG CAGATGGTTT AGTTTATGAA CCAACAGTAT TTGATCTTTC ACCACAGCAG  360
AAAGAGTGGC AGAGGATGCT GCAGCTGATT CAGAGTAGGC TACAAGAAGA GCATTCACTT  420
CAAGATGTGA TATTTAAAAG TGCTTTTAAA AGTACATCAA CAGCTCTTCC ACCAAGAGAA  480
```

```
GATGATTCAT CACAGTCTCC AAATGCATGC AGAATTCATG GCCATCTATA TGTCAATAAA   540
GTAGCAGGGA ATTTTCACAT AACAGTGGGC AAGGCAATTC CACATCCTCG TGGTCATGCA   600
CATTTGGCAG CACTTGTCAA CCATGAATCT TACAATTTTT CTCATAGAAT AGATCATTTG   660
TCTTTTGGAG AGCTTGTTCC AGCAATTATT AATCCTTTAG ATGGAACTGA AAAAATTGCT   720
ATAGATCACA ACCAGATGTT CCAATATTTT ATTACAGTTG TGCCAACAAA ACTACATACA   780
TATAAAATAT CAGCAGACAC CCATCAGTTT TCTGTGACAG AAAGGGAACG TATCATTAAC   840
CATGCTGCAG GCAGCCATGG AGTCTCTGGG ATATTTATGA AATATGATCT CAGTTCTCTT   900
ATGGTGACAG TTACTGAGGA GCACATGCCA TTCTGGCAGT TTTTTGTAAG ACTCTGTGGT   960
ATTGTTGGAG GAATCTTTTC AACAACAGGC ATGTTACATG GAATTGGAAA ATTTATAGTT  1020
GAAATAATTT GCTGTCGTTT CAGACTTGGA TCCTATAAAC CTGTCAATTC TGTTCCTTTT  1080
GAGGATGGCC ACACAGACAA CCACTTACCT CTTTTAGAAA ATAATACACA TTGA
```

A224 Protein sequence:

Gene name:	CDA14
Unigene number:	Hs.26813
Probeset Accession #:	N32912
Protein Accession #:	NP_057654
Signal sequence:	none found
Transmembrane domains:	none found
Cellular Localization:	nuclear

```
1          11         21         31         41         51
|          |          |          |          |          |
MRRLNRKKTL SLVKELDAFP KVPESYVETS ASGGTVSLIA FTIMALLTIM EFSVYQDTWM   60
KYEYEVDKDF SSKLRINIDI TVAMKCQYVG ADVLDLAETM VASADGLVYE PTVFDLSPQQ  120
KEWQRMLQLI QSRLQEEHSL QDVIFKSAFK STSTALPPRE DDSSQSPNAC RIHGHLYVNK  180
VAGNFHITVG KAIPHPRGHA HLAALVNHES YNFSHRIDHL SFGELVPAII NPLDGTEKIA  240
IDHNQMFQYF ITVVPTKLHT YKISADTHQF SVTERERIIN HAAGSRGVSG IFMKYDLSSL  300
MVTVTEEHMP FWQFFVRLCG IVGGIFSTTG MLEGIGKFIV EIICCRFRLG SYKFVNSVPF  360
EDGHTDNHLP LLRNNTH
```

Uterine

A225 DNA SEQUENCE:

Gene name:	ESTs, Weakly similar to JE0350 Anterior gradient-2 [H.sapiens]
Unigene number:	Hs.100686
Probeset Accession #:	AA487468
Nucleic Acid Accession #:	AA487468
Coding sequence:	55-555 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CGGCACCAAG AGCACTGGCC AAGTCAGCTT CTTCTGAGAG AGTCTCTAGA AGACATGATG   60
CTACACTCAG CTTTGGGTCT CTGCCTCTTA CTCGTCACAG TTTCTTCCAA CCTTGCCATT  120
GCAATAAAAA AGGAAAAGAG GCCTCCTCAG ACACTCTCAA GAGGATGGGG AGATGACATC  180
ACTTGGGTAC AAACTTATGA AGAAGGTCTC TTTTATGCTC AAAAAAGTAA GAAGCCATTA  240
ATGGTTATTC ATCACCTGGA GGATTGTCAA TACTCTCAAG CACTAAAGAA AGTATTTGCC  300
CAAAATGAAG AAATACAAGA AATGGCTCAG AATAAGTTCA TCATGCTAAA CCTTATGCAT  360
GAAACCACTG ATAAGAATTT ATCACCTGAT GGGCAATATG TGCCTAGAAT CATGTTTGTA  420
GACCCTTCTT TAACAGTTAG AGCTGACATA GCTGGAAGAT ACTCTAACAG ATTGTACACA  480
TATGAGCCTC GGGATTTACC CCTATTGATA GAAAACATGA AGAAAGCATT AAGACTTATT  540
CAGTCAGAGC TATAAGAGAT GATAGAAAAA AGCCTTCACT TCAAAGAAGT CAAATTTCAT  600
GAAGAAAACC TCTGGCACAT TGACAAATAC TAAATGTGCA AGTATATAGA TTTTGTAATA  660
TTACTATTTA GTTTTTTTAA TGTGTTTGCA ATAGTCTTAT TAAAATAAAT GTTTTTTAAA  720
TCTGAAAAAA AAAAAAAAAA AAAAAAAAA
```

A226 Protein sequence:

Gene name:	ESTs, Weakly similar to JE0350 Anterior gradient-2 [H.sapiens]
Unigene number:	Hs.100686
Probeset Accession #:	AA487468
Protein Accession #:	none found
Signal sequence:	1-23
Transmembrane domains:	none found
Cellular Localization:	secreted

```
1          11         21         31         41         51
|          |          |          |          |          |
MMLHSALGLC LLLVTVSSNL AIAIKKEKRP PQTLSRGWGD DITWVQTYEE GLFYAQKSKK   60
PLMVIHHLED CQYSQALKKV FAQNEEIQEM AQNKFIMLNL MHETTDKNLS PDGQYVPRIM  120
FVDPSLTVRA DIAGRYSNRL YTYEPRDLPL LIENMKKALR LIQSEL
```

A227 DNA SEQUENCE

Gene name:	G protein-coupled receptor 49 (GPR49) (HG38) (LGR5)
Unigene number:	Hs.285529
Probeset Accession #:	AA460530
Nucleic Acid Accession #:	NM_003667
Coding sequence:	201-2924 (underlined sequences correspond to start and stop codons)

```
         1           11          21          31          41          51
         |           |           |           |           |           |
    GTGGCGGCAA CCGGCACCTC AGTCCCCGCC GCGCTTCTCC TCGCCGCCCA CGCCGTGGGG      60
    TCAGGAACGC GGCGTCTGGC GCTGCAGACG CCCGCTGAGT TGCAGAAGCC CACGGAGCGG     120
    CGCCCGGCGC GCCACGGCCC GTAGCAGTCC GGTGCTGCTC TCCGCCCGCG TCCGGCTCGT     180
    GGCCCCCTAC TTCGGGCACC ATGGACACCT CCCGGCTCGG TGTGCTCCTG TCCTTGCCTG     240
    TGCTGCTGCA GCTGGCGACC GGGGGGCAGCT CTCCCAGGTC TGGTGTGTTG CTGAGGGGCT     300
    GCCCCACACA CTGTCATTGC GAGCCCGACG GCAGGATGTT GCTCAGGGTG GACTGCTCCG     360
    ACCTGGGGCT CTCGGGAGCTG CCTTCCAACC TCAGCGTCTT CACCTCCTAC CTAGACCTCA     420
    GTATGAACAA CATCAGTCAG CTGCTCCCGA ATCCCCTGCC CAGTCTCCGC TTCCTGGAGG     480
    AGTTACGTCT TGCGGGAAAC GCTCTGACAT ACATTCCCAA GGGAGCATTC ACTGGCCTTT     540
    ACAGTCTTAA AGTTCTTATG CTGCAGAATA ATCAGCTAAG ACACGTACCC ACAGAAGCTC     600
    TGCAGAATTT GCGAAGCCTT CAATCCCTGC GTCTGGATGC TAACCACATC AGCTATGTGC     660
    CCCCAAGCTG TTTCAGTGGC CTGCATTCCC TGAGGCACCT GTGGCTGGAT GACAATGCGT     720
    TAACAGAAAT CCCCGTCCAG GCTTTTAGAA GTTTATCGGC ATTGCAAGCC ATGACCTTGG     780
    CCCTGAACAA AATACACCAC ATACCAGACT ATGCCTTTGG AAACCTCTCC AGCTTGGTAG     840
    TTCTACATCT CCATAACAAT AGAATCCACT CCCTGGGAAA GAAATGCTTT GATGGGCTCC     900
    ACAGCCTAGA GACTTTAGAT TTAAATTACA ATAACCTTGA TGAATTCCCC ACTGCAATTA     960
    GGACACTCTC CAACCTTAAA GAACTAGGAT TTCATAGCAA CAATATCAGG TCGATACCTG    1020
    AGAAAGCATT TGTAGGCAAC CCTTCTCTTA TTACAATACA TTTCTATGAC AATCCCATCC    1080
    AATTTGTTGG GAGATCTGCT TTTCAACATT TACCTGAACT AAGAACACTG ACTCTGAATG    1140
    GTGCCTCACA AATAACTGAA TTTCCTGATT TAACTGGAAC TGCAAACCTG GAGAGTCTGA    1200
    CTTTAACTGG AGCACAGATC TCATCTCTTC CTCAAACCGT CTGCAATCAG TTACCTAATC    1260
    TCCAAGTGCT AGATCTGTCT TACAACCTAT TAGAAGATTT ACCCAGTTTT TCAGTCTGCC    1320
    AAAAGCTTCA GAAAATTGAC CTAAGACATA ATGAAATCTA CGAAATTAAA GTTGACACTT    1380
    TCCAGCAGTT GCTTAGCCTC CGATCGCTGA ATTTGGCTTG GAACAAAATT GCTATTATTC    1440
    ACCCCAATGC ATTTTCCACT TTGCCATCCC TAATAAAGCT TCCAACCTCC    1500
    TGTCGTCTTT TCCTATAACT GGGTTACATG GTTTAACTCA CTTAAAATTA ACAGGAAATC    1560
    ATGCCTTACA GAGCTTGATA TCATCTGAAA ACTTTCCAGA ACTCAAGGTT ATAGAAATGC    1620
    CTTATGCTTA CCAGTGCTGT GCATTTGGAG TGTGTGAGAA TGCCTATAAG ATTTCTAATC    1680
    AATGGAATAA AGGTGACAAC AGCAGTATGG ACGACCTTCA TAAGGAAAGAT GCTGGAATGT    1740
    TTCAGGCTCA AGATGAACGT GACCTTGAAG ATTTCCTGCT TGACTTTGAG GAAGACCTGA    1800
    AAGCCCTTCA TTCAGTGCAG TGTTCACCTT CCCCAGGCCC CTTCAAACCC TGTGAACACC    1860
    TGCTTGATGG CTGGCTGATC AGAATTGGAG TGTGGACCAT AGCAGTTCTG GCACTTACTT    1920
    GTAATGCTTT GGTGACTTCA ACAGTTTTCA GATCCCCTCT GTACATTTCC CCCATTAAAC    1980
    TGTTAATTGG GGTCATCGCA GCAGTGAACA TGCTCACGGG AGTCTCCAGT GCCGTGCTGG    2040
    CTGGTGTGGA TGCGTTCACT TTTGGCAGCT TTGCACGACA TGGTGCCTGG TGGGAGAATG    2100
    GGGTTGGTTG CCATGTCATT GGTTTTTTGT CCATTTTTGC TTCAGAATCA TCTGTTTTCC    2160
    TGCTTACTCT GGCAGCCCTG GAGCGTGGGT TCTCTGTGAA ATATTCTGCA AAATTTGAAA    2220
    CGAAAGCTCC ATTTTCTAGC CTGAAAGTAA TCATTTTGCT CTGTGCCCTG CTGGCCTTGA    2280
    CCATGGCCGC AGTTCCCCTG CTGGGTGGCA GCAAGTATGG CGGCTCCCCT CTCTGCCTGC    2340
    CTTTGCCTTT TGGGGAGCCC AGCACCATGG GCTACATGGT CGCTCTCATC TTGCTCAATT    2400
    CCCTTTGCTT CCTCATGATG ACCATTGCCT ACACCAAGCT CTACTGCAAT TTGGACAAGG    2460
    GAGACCTGGA GAATATTTGG GACTGCTCTA TGGTAAAACA CATTGCCCTG TTGCTCTTCA    2520
    CCAACTGCAT CCTAAACTGC CCTGTGGCTT TCTTGTCCTT CTCCTCTTTA ATAAACCTTA    2580
    CATTTATCAG TCCTGAAGTA ATTAAGTTTA TCCTTCGGT GGTAGTCCCA CTTCCTGCACC    2640
    GTCTCAATCC CCTTCTCTAC ATCTTGTTCA ATCCTCACTT TAAGGAGGAT CTGGTGAGCC    2700
    TGAGAAAGCA AACCTACGTC TGGACAAGAT CAAAACACCC AAGCTTGATG TCAATTAACT    2760
    CTGATGATGT CGAAAAACAG TCCTGTGACT CAACTCAAGC CTTGGTAACC TTTACCAGCT    2820
    CCAGCATCAC TTATGACCTG CCTCCCCAGTT CCGTGCCATC ACCAGCTTAT CCAGTGACTG    2880
    AGAGCTGCCA TCTTTCCTCT GTGGCATTTG TCCCATGTCT CTAATTAATA TGTGAAGGAA    2940
    AATGTTTTCA AAGGTTGAGA ACCTGAAAAT GTGAGATTGA GTATATCAGA GCAGTAATTA    3000
    ATAAGAAGAG CTGAGGTGAA ACTCGGTTTA AA
```

A228 Protein sequence
Gene name: G protein-coupled receptor 49 (GPR49) (HG38) (LGR5)
Unigene number: Hs.285529
Protein Accession #:   NP_003658.1
Signal sequence: 1-22
Transmembrane domains: 557-579, 596-618, 641-663, 687-709, 724-746, 770-792, 803-825
Cellular Localization: plasma membrane

```
         1           11          21          31          41          51
         |           |           |           |           |           |
    MDTSRLGVLL SLPVLLQLAT GGSSPRSGVL LRGCPTHCHC EPDGRMLLRV DCSDLGLSEL      60
    PSNLSVFTSY LDLSMNNISQ LLPNPLPSLR FLEELRLAGN ALTYIPKGAF TGLYSLKVLM     120
    LQNNQLRHVP TEALQNLRSL QSLRLDANHI SYVPPSCFSG LHSLRHLWLD DNALTEIPVQ     180
    AFRSLSALQA MTLALNKIHH IPDYAFGNLS SLVVLHLHNN RIHSLGKKCF DGLHSLETLD     240
    LNYNNLDEPP TAIRTLSNLK ELGFHSNNIR SIPEKAFVGN PSLITIHFYD NPIQFVGRSA     300
    FQHLPELRTL TLNGASQITE FPDLTGTANL ESLTLTGAQI SSLPQTVCNQ LPNLQVLDLS     360
    YNLLEDLPSF SVCQKLQKID LRHNEIYEIK VDTFQQLLSL RSLNLAWNKI AIIHPNAFST     420
    LPSLIKLDLS SNLLSSFPIT GLHGLTHLKL TGNHALQSLI SSENFPELKV IEMPYAYQCC     480
    AFGVCENAYK ISNQWNKGDN SSMDDLHKED AGMFQAQDER DLEDFLIDPE EDLXALHSVQ     540
    CSPSPGPFKP CEHLLDGWLI RIGVWTIAVL ALTCNALVTS TVFRSPLYIS PIKLLIGVIA     600
    AVNMLTGVSS AVLAGVDAFT FGSFARHGAW WENGVGCHVI GFLSIFASES SVFLLTLAAL     660
    ERGFSVKYSA KFETKAPFSS LKVIILLCAL LALTMAAVPL LGGSKYGASP LCLPLPFGEP     720
    STMGYMVALI LLNSLCFLMM TIAYTKLYCN LDKGDLENIW DCSMVKHIAL LLFTNCILNC     780
    FVAFLSFSSL INLTFISPEV IKFILLVVVP LPACLNPLLY ILFNPHFKED LVSLRKQTYV     840
    WTRSKHPSLM SINSDDVEKQ SCDSTQALVT FTSSSITYDL PPSSVPSPAY PVTESCHLSS     900
    VAFVPCL
```

Table 75: See Table 1

Table 76A depicts Seq ID No; UnigeneID; UnigeneTitle; PKey; Predicted Cellular Localization; and Exemplar Accession for each of the sequences in Table 76. The information in Table 76A is linked by SeqID No to Table 76.

Seq ID No:	Sequence ID No for sequences in table
Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
Pred Subcell Loc:	Predicted sub-cellular localization

Seq ID No	Pkey	ExAccn	UnigeneID	Unigene Title	Pred Subcell Loc
Seq ID 1 & 2	425023	AW956889	Hs.154210	endothelial differentiation, sphingolipi	plasma membrane
Seq ID 3 & 4	424503	NM_002205	Hs.149609	Integrin, alpha 5 (fibronectin receptor,	plasma membrane
Seq ID 5 & 6	429423	AI016712	Hs.287797	Integrin, beta 1 (fibronectin receptor,	plasma membrane
Seq ID 7 & 8	400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	secreted
Seq ID 9 & 10	419172	AW338625	Hs.22120	ESTs; similar to TRANSMEMBRANE 4 SUPERF	plasma membrane
Seq ID 11 & 12	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	secreted
Seq ID 13 & 14	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	secreted
Seq ID 15 & 16	407836	T79340	Hs.200272	B-cell CLL/lymphoma 6, member B, zinc fi	intracell
Seq ID 17 & 18	414577	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to	secreted
Seq ID 19 & 20	418738	AW388633	Hs.6562	solute carrier family 7, (cationic amino	plasma membrane
Seq ID 21 & 22	428368	BE440042	Hs.63326	matrix metalloproteinase 3 (stromelysin	secreted
Seq ID 23 & 24	415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	secreted
Seq ID 25 & 26	429276	AF056085	Hs.198612	G protein-coupled receptor 51	plasma membrane
Seq ID 27 & 28	418994	AA296520	Hs.89546	selectin E (endothelial adhesion molecul	plasma membrane
Seq ID 29 & 30	407975	X89426	Hs.41716	endothelial cell-specific molecule 1	secreted
Seq ID 31 & 32	429113	D28235	Hs.196384	prostaglandin-endoperoxide synthase 2 (p	intracell
Seq ID 33 & 34	418506	AA084248	Hs.85339	G protein-coupled receptor 39	plasma membrane
Seq ID 35 & 36	423961	D13666	Hs.136348	periostin (OSF-2os)	secreted
Seq ID 37 & 38	414812	X72755	Hs.77367	monokine induced by gamma interferon	secreted
Seq ID 39 & 40	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	plasma membrane
Seq ID 41 & 42	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	plasma membrane
Seq ID 43 & 44	424399	AI905687		AI905687:IL-BT095-190199-019 BT095 Homo	secreted
Seq ID 45 & 46	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	secreted
Seq ID 47 & 48	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy	secreted
Seq ID 49 & 50	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti	secreted
Seq ID 51 & 52	439569	AW602166	Hs.222399	CEGP1 protein	secreted
Seq ID 53 & 54	411558	AA102570	Hs.70725	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
Seq ID 55 & 56	400303	AA242756	Hs.79136	LIV-1 protein, estrogen regulated	plasma membrane
Seq ID 57 & 58	411789	AF245505	Hs.72157	Adlican	secreted
Seq ID 59 & 60	428698	AA852773	Hs.334838	KIAA1866 protein	plasma membrane
Seq ID 61 & 62	450098	W27249	Hs.8109	hypothetical protein FLJ21080	intracell
Seq ID 63 & 64	421552	AF026692	Hs.105700	secreted frizzled-related protein 4	secreted
Seq ID 65 & 66	452747	BE153855	Hs.61460	Ig superfamily receptor LNIR	plasma membrane
Seq ID 67 & 68	415539	AI733881	Hs.72472	BMP-R1B	plasma membrane
Seq ID 69 & 70	416636	N32536	Hs.42645	solute carrier family 16 (monocarboxylic	plasma membrane
Seq ID 71 & 72	416636	N32536	Hs.42645	solute carrier family 16 (monocarboxylic	plasma membrane
Seq ID 73 & 74	409079	W87707	Hs.82065	interleukin 6 signal transducer (gp130,	plasma membrane
Seq ID 75 & 76	442082	R41823	Hs.7413	ESTs	plasma membrane
Seq ID 77 & 78	400297	AI127076	Hs.306201	hypothetical protein DKFZp564O1278	plasma membrane
Seq ID 79 & 80	451398	AI793124	Hs.144479	ESTs	intracell
Seq ID 81 & 82	429220	AW207206		ESTs	plasma membrane
Seq ID 83 & 84	421524	AA312082	Hs.105445	GDNF family receptor alpha 1	plasma membrane
Seq ID 85 & 86	423242	AL039402	Hs.125783	DEME-6 protein	plasma membrane
Seq ID 87 & 88	423242	AL039402	Hs.125783	DEME-6 protein	intracell
Seq ID 89 & 90	423242	AL039402	Hs.125783	DEME-6 protein	intracell
Seq ID 91 & 92	452190	H26735	Hs.91668	Homo sapiens clone PP1498 unknown mRNA	intracell
Seq ID 93 & 94	452190	H26735	Hs.91668	Homo sapiens clone PP1498 unknown mRNA	intracell
Seq ID 95 & 96	325372			Phase 2 & 3 Exons	cytoplasmic
Seq ID 97 & 98	450375	AA009847		a disintegrin and metalloproteinase doma	plasma membrane
Seq ID 99 & 100	426215	AW963419	Hs.155223	stanniocalcin 2	secreted
Seq ID 101 & 102	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	secreted
Seq ID 103 & 104	429353	AL117406	Hs.335891	ATP-binding cassette transporter MRP8	plasma membrane
Seq ID 105 & 106	429353	AL117406	Hs.335891	ATP-binding cassette transporter MRP8	plasma membrane
Seq ID 107 & 108	432201	AI538613	Hs.298241	Transmembrane protease, serine 3	plasma membrane
Seq ID 109 & 110	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	secreted
Seq ID 111 & 112	446163	AA026880	Hs.25252	prolactin receptor	plasma membrane
Seq ID 113 & 114	442117	AW664964	Hs.128899	ESTs; hypothetical protein for IMAGE:447	plasma membrane
Seq ID 115 & 116	428179	AI127772	Hs.279696	serum/glucocorticoid regulated kinase-l	intracell
Seq ID 117 & 118	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	plasma membrane
Seq ID 119 & 120	447033	AI357412	Hs.157601	ESTs	secreted
Seq ID 121 & 122	447033	AI357412	Hs.157601	ESTs	secreted
Seq ID 123 & 124	447033	AI357412	Hs.157601	ESTs	secreted
Seq ID 125 & 126	115522	BE614387	Hs.333893	c-Myc target JPO1	intracell
Seq ID 127 & 128	452679	Z42387	Hs.83883	transmembrane, prostate androgen induced	plasma membrane
Seq ID 129 & 130	446051	BE048061	Hs.37054	ephrin-A3	plasma membrane
Seq ID 131 & 132	422048	NM_012445	Hs.288126	spondin 2, extracellular matrix protein	secreted
Seq ID 133 & 134	410418	D31382	Hs.63325	transmembrane protease, serine 4	plasma membrane
Seq ID 135 & 136	446342	BE298665	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr	plasma membrane
Seq ID 137 & 138	422260	AA315993	Hs.105484	regenerating gene type IV	secreted

Seq ID					
Seq ID 139 & 140	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	secreted
Seq ID 141 & 142	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	secreted
Seq ID 143 & 144	111929	AF027208	Hs.112360	prominin (mouse)-like 1	plasma membrane
Seq ID 145 & 146	111929	AF027208	Hs.112360	prominin (mouse)-like 1	plasma membrane
Seq ID 147 & 148	111929	AF027208	Hs.112360	prominin (mouse)-like 1	plasma membrane
Seq ID 149 & 150	104888	AW939591	Hs.5940	mucin 13, epithelial transmembrane	plasma membrane
Seq ID 151 & 152	420159	AI572490	Hs.99785	Homo sapiens cDNA: FLJ21245 fis, clone C	plasma membrane
Seq ID 153 & 154	422330	D30783	Hs.115263	epiregulin	plasma membrane
Seq ID 155 & 156	452461	N78223	Hs.108106	transcription factor	intracell
Seq ID 157 & 158	413324	V00571	Hs.75294	corticotropin releasing hormone	secreted
Seq ID 159 & 160	412420	AL035668	Hs.73853	bone morphogenetic protein 2	secreted
Seq ID 161 & 162	416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	secreted
Seq ID 163 & 164	407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	secreted
Seq ID 165	402230			Fgenesh predicted: CYTOCHROME P450 4F5 (	intracell
Seq ID 166 & 167	402230			Fgenesh predicted: CYTOCHROME P450 4F5 (	intracell
Seq ID 168 & 169	432829	W60377	Hs.57772	ESTs	intracell
Seq ID 170 & 171	425721	AC002115	Hs.159309	uroplakin 1A	plasma membrane
Seq ID 172 & 173	420370	Y13645	Hs.97234	uroplakin 2	plasma membrane
Seq ID 174 & 175	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365O12.1 [H.sa	plasma membrane
Seq ID 176 & 177	402075			ENSP00000251056*;Plasma membrane calcium	secreted
Seq ID 178 & 179	421110	AJ250717	Hs.1355	cathepsin E	secreted
Seq ID 180 & 181	451668	Z43948	Hs.326444	cartilage acidic protein 1	plasma membrane
Seq ID 182 & 183	451668	Z43948	Hs.326444	cartilage acidic protein 1	secreted
Seq ID 184 & 185	451668	Z43948	Hs.326444	cartilage acidic protein 1	intracell
Seq ID 186 & 187	408243	Y00787	Hs.624	interleukin 8	secreted
Seq ID 188 & 189	422282	AF019225	Hs.114309	apolipoprotein L	secreted
Seq ID 190 & 191	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member	plasma membrane
Seq ID 192 & 193	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	plasma membrane
Seq ID 194 & 195	404875			NM_022819*:Homo sapiens phospholipase A2	intracell
Seq ID 196 & 197	425883	AL137708	Hs.161031	Homo sapiens mRNA; cDNA DKFZp434K0322 (f	plasma membrane
Seq ID 198 & 199	404977			insulin-like growth factor 2 (somatomed	secreted
Seq ID 200 & 201	420876	AA918425	Hs.177744	ESTs	plasma membrane
Seq ID 202 & 203	427747	AW411425	Hs.180655	serine/threonine kinase 12	intracell
Seq ID 204 & 205	420281	AI623693	Hs.323494	Predicted cation efflux pump	plasma membrane
Seq ID 206 & 207	446673	NM_018361	Hs.15871	LPAP for lysophosphatidic acid phosphata	intracell
Seq ID 208	437553	AI829935	Hs.130497	ESTs, Weakly similar to MATB_HUMAN CHLOR	plasma membrane
Seq ID 209 & 210	437553	AI829935	Hs.130497	ESTs, Weakly similar to MATB_HUMAN CHLOR	plasma membrane
Seq ID 211 & 212	437553	AI829935	Hs.130497	ESTs, Weakly similar to MATB_HUMAN CHLOR	plasma membrane
Seq ID 213 & 214	426900	AW163564	Hs.142375	ESTs	plasma membrane
Seq ID 215 & 216	426900	AW163564	Hs.142375	ESTs	plasma membrane
Seq ID 217 & 218	426900	AW163564	Hs.142375	ESTs	plasma membrane
Seq ID 219 & 220	405932			C15000305;gi|3806122|gb|AAC69198.1| (AF0	intracell
Seq ID 221 & 222	405932			C15000305;gi|3806122|gb|AAC69198.1| (AF0	intracell
Seq ID 223 & 224	424008	R02740	Hs.137555	putative chemokine receptor; GTP-binding	plasma membrane
Seq ID 225 & 226	424008	R02740	Hs.137555	putative chemokine receptor; GTP-binding	plasma membrane
Seq ID 227 & 228	424008	R02740	Hs.137555	putative chemokine receptor; GTP-binding	plasma membrane
Seq ID 229 & 230	444342	NM_014354	Hs.10887	similar to lysosome-associated membrane	secreted
Seq ID 231 & 232	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy	secreted
Seq ID 233 & 234	417079	U65590	Hs.81134	interleukin 1 receptor antagonist	secreted
Seq ID 235 & 236	412986	X81120	Hs.75110	cannabinoid receptor 1 (brain)	plasma membrane
Seq ID 237 & 238	412986	X81120	Hs.75110	cannabinoid receptor 1 (brain)	plasma membrane
Seq ID 239 & 240	412986	X81120	Hs.75110	cannabinoid receptor 1 (brain)	plasma membrane
Seq ID 241 & 242	447072	D61594	Hs.17279	tyrosylprotein sulfotransferase 1	intracell
Seq ID 243 & 244	419723	AL120193	Hs.339810	longevity assurance (LAG1, S. cerevisiae	secreted
Seq ID 245 & 246	419723	AL120193	Hs.339810	longevity assurance (LAG1, S. cerevisiae	plasma membrane
Seq ID 247 & 248	430890	X54232	Hs.2699	glypican 1	plasma membrane
Seq ID 249 & 250	456759	BE269150	Hs.127792	delta (Drosophila)-like 3	plasma membrane
Seq ID 251	429466	M85835	Hs.12827	ESTs	
Seq ID 252	429466	M85835	Hs.12827	ESTs	
Seq ID 253 & 254	419721	NM_001650	Hs.288650	aquaporin 4	plasma membrane
Seq ID 255 & 256	407034	U84540		gb:Human dystrobrevin isoform DTN-3 (DTN	secreted
Seq ID 257 & 258	413472	BE242870	Hs.75379	solute carrier family 1 (glial high affi	plasma membrane
Seq ID 259 & 260	436380	T05430	Hs.6194	chondroitin sulfate proteoglycan BEHAB/b	secreted & plasma membrane
Seq ID 261 & 262	426271	AF026547	Hs.169047	chondroitin sulfate proteoglycan 3 (neur	secreted
Seq ID 263 & 264	419704	AA429104	Hs.45057	ESTs	intracell
Seq ID 265 & 266	444471	AB020684	Hs.11217	KIAA0877 protein	plasma membrane
Seq ID 267 & 268	409395	U46745	Hs.336678	dystrobrevin, alpha	secreted
Seq ID 269 & 270	413063	AL035737	Hs.75184	chitinase 3-like 1 (cartilage glycoprote	secreted
Seq ID 271 & 272	433800	AI034361	Hs.135150	lung type-I cell membrane-associated gly	plasma membrane
Seq ID 273	458435	AI416718	Hs.144121	ESTs, Weakly similar to T46916 hypotheti	
Seq ID 274	458435	AI416718	Hs.144121	ESTs, Weakly similar to T46916 hypotheti	
Seq ID 275 & 276	424343	AW956360	Hs.4748	adenylate cyclase activating polypeptide	plasma membrane
Seq ID 277 & 278	424998	U58515	Hs.154138	chitinase 3-like 2	secreted
Seq ID 279 & 280	412709	AL022327	Hs.74518	KIAA0027 protein	plasma membrane
Seq ID 281 & 282	435615	Y15065	Hs.4975	potassium voltage-gated channel, KQT-lik	plasma membrane
Seq ID 283 & 284	404049			NM_018937*:Homo sapiens protocadherin be	plasma membrane
Seq ID 285 & 286	418932	L34059	Hs.89484	cadherin 4, type 1, R-cadherin (retinal)	plasma membrane
Seq ID 287 & 288	404029			NM_018936*:Homo sapiens protocadherin be	plasma membrane
Seq ID 289 & 290	436480	AJ271643	Hs.87469	putative acid-sensing ion channel	intracell
Seq ID 291 & 292	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro	secreted
Seq ID 293 & 294	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro	secreted
Seq ID 295 & 296	436895	AF037335	Hs.5338	carbonic anhydrase XII	plasma membrane
Seq ID 297 & 298	421471	U90545	Hs.327179	solute carrier family 17 (sodium phospha	plasma membrane

Seq ID				Description	Location
Seq ID 299 & 300	428296	NM_003058	Hs.183572	solute carrier family 22 (organic cation	plasma membrane
Seq ID 301 & 302	423508	AW604297	Hs.129711	hepatitis A virus cellular receptor 1	plasma membrane
Seq ID 303 & 304	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitte	plasma membrane
Seq ID 305 & 306	410407	X66839	Hs.63287	carbonic anhydrase IX	plasma membrane
Seq ID 307 & 308	453496	AA442103	Hs.33084	solute carrier family 2 (facilitated glu	plasma membrane
Seq ID 309 & 310	420737	L08096	Hs.99899	CD70 ; tumor necrosis factor (ligand) s	plasma membrane
Seq ID 311 & 312	309931	AW341683		gb:hd13d01.x1 Soares_NFL_T_GBC_S1 Homo s	plasma membrane
Seq ID 313 & 314	412719	AW016610	Hs.816	ESTs	intracell
Seq ID 315 & 316	417034	NM_006183	Hs.80962	neurotensin	secreted
Seq ID 317 & 318	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	plasma membrane
Seq ID 319 & 320	413753	U17760	Hs.75517	laminin, beta 3 (nicein (125kD), kalinin	secreted
Seq ID 321 & 322	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	plasma membrane
Seq ID 323 & 324	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	secreted
Seq ID 325 & 326	418663	AK001100	Hs.41690	desmocollin 3	plasma membrane
Seq ID 327 & 328	418663	AK001100	Hs.41690	desmocollin 3	plasma membrane
Seq ID 329 & 330	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	secreted
Seq ID 331 & 332	406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad	plasma membrane
Seq ID 333 & 334	431846	BE019924	Hs.271580	uroplakin 1B	plasma membrane
Seq ID 335 & 336	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	secreted
Seq ID 337 & 338	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	plasma membrane
Seq ID 339 & 340	437044	AL035864	Hs.69517	differentially expressed in Fanconi's en	plasma membrane
Seq ID 341 & 342	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino	plasma membrane
Seq ID 343 & 344	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	plasma membrane
Seq ID 345 & 346	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor	secreted
Seq ID 347 & 348	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	plasma membrane
Seq ID 349 & 350	417542	J04129	Hs.82269	progestagen-associated endometrial prote	secreted
Seq ID 351 & 352	449230	BE613348	Hs.211579	melanoma cell adhesion molecule	plasma membrane
Seq ID 353 & 354	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma	plasma membrane
Seq ID 355 & 356	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma	plasma membrane
Seq ID 357 & 358	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	secreted
Seq ID 359 & 360	418462	BE001596	Hs.85266	integrin, beta 4	plasma membrane
Seq ID 361 & 362	410274	AA381807	Hs.61762	hypoxia-inducible protein 2	secreted
Seq ID 363 & 364	439606	W79123	Hs.58561	G protein-coupled receptor 87	plasma membrane
Seq ID 365 & 366	404877			NM_005365:Homo sapiens melanoma antigen,	intracell
Seq ID 367 & 368	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote	plasma membrane
Seq ID 369 & 370	109424	NM_005329	Hs.85962	hyaluronan synthase 3	plasma membrane
Seq ID 371 & 372	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
Seq ID 373 & 374	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
Seq ID 375 & 376	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
Seq ID 377 & 378	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
Seq ID 379 & 380	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
Seq ID 381 & 382	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
Seq ID 383 & 384	421817	AF146074	Hs.106560	ATP-binding cassette, sub-family C (CFTR	plasma membrane
Seq ID 385 & 386	418678	NM_001327	Hs.87225	cancer/testis antigen (NY-ESO-1)	intracell
Seq ID 387 & 388	418678	NM_001327	Hs.87225	cancer/testis antigen (NY-ESO-1)	intracell
Seq ID 389 & 390	409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini	secreted
Seq ID 391 & 392	332180	AF134160	Hs.7327	claudin 1	plasma membrane
Seq ID 393 & 394	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	plasma membrane
Seq ID 395 & 396	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	plasma membrane
Seq ID 397 & 398	439223	AW238299	Hs.250618	UL16 binding protein 2	plasma membrane
Seq ID 399 & 400	409757	NM_001898	Hs.123114	cystatin SN	secreted
Seq ID 401 & 402	428969	AF120274	Hs.194689	artemin	secreted
Seq ID 403 & 404	428969	AF120274	Hs.194689	artemin	secreted
Seq ID 405 & 406	428969	AF120274	Hs.194689	artemin	secreted
Seq ID 407 & 408	428969	AF120274	Hs.194689	artemin	secreted
Seq ID 409 & 410	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-	secreted
Seq ID 411 & 412	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-	secreted
Seq ID 413 & 414	414774	X02419	Hs.77274	plasminogen activator, urokinase	secreted
Seq ID 415 & 416	407944	R34008	Hs.239727	desmocollin 2	plasma membrane
Seq ID 417 & 418	407944	R34008	Hs.239727	desmocollin 2	plasma membrane
Seq ID 419 & 420	428486	AW583497	Hs.184604	pancreatic polypeptide	secreted
Seq ID 421 & 422	457489	AI693815	Hs.127179	cryptic gene	secreted
Seq ID 423 & 424	432874	W94322	Hs.279651	melanoma inhibitory activity	secreted
Seq ID 425 & 426	445891	AW391342	Hs.199460	DPCR1 protein	plasma membrane
Seq ID 427 & 428	445891	AW391342	Hs.199460	DPCR1 protein	plasma membrane
Seq ID 429 & 430	404682			C9001138*:gi|12738842|ref|NP_073725.1| p	secreted
Seq ID 431 & 432	429547	AW009166	Hs.99376	ESTs	secreted
Seq ID 433 & 434	425921	NM_007231	Hs.162211	solute carrier family 6 (neurotransmitte	plasma membrane
Seq ID 435 & 436	407242	M18728		gb:Human nonspecific crossreacting antig	plasma membrane
Seq ID 437 & 438	407242	M18728		gb:Human nonspecific crossreacting antig	plasma membrane
Seq ID 439 & 440	407242	M18728		gb:Human nonspecific crossreacting antig	plasma membrane
Seq ID 441 & 442	432598	AJ224741	Hs.278461	matrilin 3	secreted
Seq ID 443 & 444	444006	BE395085	Hs.10086	type I transmembrane protein Fn14	plasma membrane
Seq ID 445 & 446	423685	BE350494	Hs.49753	uveal autoantigen with coiled coil domai	intracell
Seq ID 447 & 448	428392	H10233	Hs.2265	secretory granule, neuroendocrine protei	secreted
Seq ID 449 & 450	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	plasma membrane
Seq ID 451 & 452	446030	N30714	Hs.325960	membrane-spanning 4-domains, subfamily A	plasma membrane
Seq ID 453 & 454	422109	S73265	Hs.1473	gastrin-releasing peptide	secreted
Seq ID 455 & 456	419235	AW470411	Hs.288433	neurotrimin	plasma membrane
Seq ID 457 & 458	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc	plasma membrane
Seq ID 459 & 460	427333	AF087797	Hs.176658	aquaporin 8	plasma membrane
Seq ID 461 & 462	417931	W95642	Hs.82961	trefoil factor 3 (intestinal)	secreted
Seq ID 463 & 464	419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy	secreted

Seq ID 465 & 466	431629	AU077025	Hs.265827	Interferon, alpha-inducible protein (clo	secreted
Seq ID 467 & 468	413554	AA319146	Hs.75426	secretogranin II (chromogranin C)	secreted
Seq ID 469 & 470	452194	AI694413	Hs.332649	Ubiquitin-like protein FAT10???	plasma membrane
Seq ID 471 & 472	452194	AI694413	Hs.332649	Ubiquitin-like protein FAT10???	plasma membrane
Seq ID 473 & 474	426322	J05068	Hs.2012	transcobalamin I (vitamin B12 binding pr	secreted
Seq ID 475 & 476	429010	Y18198	Hs.194725	one cut domain, family member 2	intracell
Seq ID 477 & 478	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu	secreted
Seq ID 479 & 480	448243	AW369771	Hs.52620	integrin, beta 8	plasma membrane
Seq ID 481 & 482	426427	M86699	Hs.169840	TTK protein kinase	intracell
Seq ID 483 & 484	428187	AI687303	Hs.285529	G protein-coupled receptor 49	plasma membrane
Seq ID 485 & 486	428579	NM_005756	Hs.184942	G protein-coupled receptor 64	plasma membrane
Seq ID 487 & 488	428579	NM_005756	Hs.184942	G protein-coupled receptor 64	plasma membrane
Seq ID 489 & 490	428579	NM_005756	Hs.184942	G protein-coupled receptor 64	plasma membrane
Seq ID 491 & 492	428579	NM_005756	Hs.184942	G protein-coupled receptor 64	plasma membrane
Seq ID 493 & 494	445537	AJ245671	Hs.12844	EGF-like-domain, multiple 6	secreted
Seq ID 495 & 496	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 6	plasma membrane
Seq ID 497 & 498	424620	AA101043	Hs.151254	kallikrein 7 (chymotryptic, stratum corn	secreted
Seq ID 499 & 500	406400			kallikrein 8 (neuropsin/ovasin) (KLK8)	secreted
Seq ID 501 & 502	431130	NM_006103	Hs.2719	HE4; epididymis-specific, whey-acidic pr	secreted
Seq ID 503 & 504	420440	NM_002407	Hs.97644	mammaglobin 2	secreted
Seq ID 505 & 506	428450	NM_014791	Hs.184339	KIAA0175 gene product	intracell
Seq ID 507 & 508	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	secreted
Seq ID 509 & 510	436982	AB018305	Hs.5378	spondin 1, (f-spondin) extracellular mat	secreted
Seq ID 511 & 512	453392	U23752	Hs.32964	SRY (sex determining region Y)-box 11	intracell
Seq ID 513	431989	AW972870	Hs.291069	ESTs	
Seq ID 514	439820	AL360204	Hs.283653	Homo sapiens mRNA full length insert cDN	
Seq ID 515 & 516	409178	BE393948	Hs.50915	kallikrein 5	secreted
Seq ID 517 & 518	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic	secreted
Seq ID 519 & 520	421478	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos	plasma membrane
Seq ID 521 & 522	421478	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos	plasma membrane
Seq ID 523 & 524	425776	U25128	Hs.159499	parathyroid hormone receptor 2	plasma membrane
Seq ID 525 & 526	425776	U25128	Hs.159499	parathyroid hormone receptor 2	plasma membrane
Seq ID 527 & 528	452097	AB002364	Hs.27916	a disintegrin-like and metalloprotease (	secreted
Seq ID 529 & 530	416530	U62801	Hs.79361	kallikrein 6 (neurosin, zyme)	secreted
Seq ID 531 & 532	431515	NM_012152	Hs.256583	endothelial differentiation, lysophospha	plasma membrane
Seq ID 533 & 534	419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	plasma membrane
Seq ID 535 & 536	412078	X69699	Hs.73149	paired box gene 8	intracell
Seq ID 537 & 538	412078	X69699	Hs.73149	paired box gene 8	intracell
Seq ID 539 & 540	411773	NM_006799	Hs.72026	protease, serine, 21 (testisin)	secreted
Seq ID 541 & 542	407792	AI077715	Hs.39384	putative secreted ligand homologous to f	secreted
Seq ID 543 & 544	431616	AA508552	Hs.195839	ESTs, Weakly similar to I38022 hypotheti	plasma membrane
Seq ID 545 & 546	452792	AB037765	Hs.30852	KIAA1344 protein	plasma membrane
Seq ID 547 & 548	400294	N95796	Hs.278695	Homo sapiens prostein mRNA, complete cds	plasma membrane
Seq ID 549 & 550	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	plasma membrane
Seq ID 551 & 552	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	plasma membrane
Seq ID 553 & 554	432853	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	plasma membrane
Seq ID 555 & 556	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	plasma membrane
Seq ID 557 & 558	400290	H18836	Hs.31608	hypothetical protein FLJ20041	plasma membrane
Seq ID 559 & 560	410001	AB041036	Hs.57771	kallikrein 11	secreted
Seq ID 561 & 562	418396	AI765805	Hs.26691	ESTs	plasma membrane
Seq ID 563 & 564	451027	AW519204	Hs.40808	ESTs	plasma membrane
Seq ID 565 & 566	446057	AI120227	Hs.149358	ESTs, Weakly similar to A46010 X-linked	plasma membrane
Seq ID 567 & 568	433466	AA508353	Hs.105314	relaxin 1 (H1)	secreted
Seq ID 569 & 570	453370	AI470523	Hs.139336	ATP-binding cassette, sub-family C (CFTR	plasma membrane
Seq ID 571 & 572	453370	AI470523	Hs.139336	ATP-binding cassette, sub-family C (CFTR	plasma membrane
Seq ID 573 & 574	414569	AF109298	Hs.118258	prostate cancer associated protein 1	plasma membrane
Seq ID 575 & 576	413435	X51405	Hs.75360	carboxypeptidase E	secreted
Seq ID 577 & 578	426501	AW043782	Hs.293818	ESTs	plasma membrane
Seq ID 579 & 580	448099	AF179274	Hs.22791	transmembrane protein with EGF-like and	plasma membrane
Seq ID 581 & 582	408369	R38438	Hs.182575	solute carrier family 15 (H??? transport	plasma membrane
Seq ID 583 & 584	412828	AI972402	Hs.306051	hypothetical protein MGC2648	secreted
Seq ID 585 & 586	403047			NM_005656*:Homo sapiens transmembrane pr	plasma membrane
Seq ID 587 & 588	403047			NM_005656*:Homo sapiens transmembrane pr	plasma membrane
Seq ID 589 & 590	408430	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine	plasma membrane
Seq ID 591 & 592	445413	AA151342	Hs.12677	CGI-147 protein	secreted
Seq ID 593 & 594	451982	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f	intracell
Seq ID 595 & 596	427958	AA418000	Hs.98280	potassium intermediate/small conductance	plasma membrane
Seq ID 597 & 598	421887	AW161450	Hs.109201	CGI-66 protein	plasma membrane
Seq ID 599 & 600	425071	NM_013989	Hs.154424	deiodinase, iodothyronine, type II	secreted
Seq ID 601 & 602	432101	AI918950	Hs.123642	EphA3	plasma membrane
Seq ID 603 & 604	407786	AA687538	Hs.38972	tetraspan 1	plasma membrane
Seq ID 605 & 606	416836	D54745	Hs.80247	cholecystokinin	secreted
Seq ID 607 & 608	416539	Y07909	Hs.79368	epithelial membrane protein 1	plasma membrane
Seq ID 609 & 610	131083	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
Seq ID 611 & 612	131083	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
Seq ID 613 & 614	131083	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
Seq ID 615 & 616	131083	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
Seq ID 617 & 618	422424	AI186431	Hs.295638	prostate differentiation factor	secreted
Seq ID 619 & 620	428970	BE276891	Hs.194691	retinoic acid induced 3	plasma membrane
Seq ID 621 & 622	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	secreted
Seq ID 623 & 624	439018	AW300887	Hs.26638	membrane-spanning 4-domains, subfamily A	plasma membrane
Seq ID 625 & 626	420610	AI683183	Hs.99348	distal-less homeo box 5	intracell
Seq ID 627 & 628	425723	NM_014420	Hs.159311	dickkopf (Xenopus laevis) homolog 4	secreted

Seq ID 629 & 630	456662	NM_002448	Hs.1494	msh (Drosophila) homeo box homolog 1 (fo	intracell
Seq ID 631 & 632	418281	U09550	Hs.1154	oviductal glycoprotein 1, 120kD (mucin 9	secreted
Seq ID 633 & 634	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	intracell

Table 76B:

Pkey:	Unique Eos probeset identifier number
CAT number:	Gene cluster number
Accession:	Genbank accession numbers

Pkey	CAT Number	Accession
424399	238961_1	AI905687 AI905624 AI905837 AI905623 AA340069 R75793 W72837 BE074512 AI905633 W72838 BE092421 AI127172 BE186013 AW070916 AI139456 AW176044 AW291950
429220	301384_1	AW207206 AW341473 AA448195 AI951341
450375	83327_1	AA009647 AA131254 AA374293 AW954405 H04410 AW606284 AA151166 BE157467 BE157601 H04384 W46291 AW663674 H04021 H01532 AA190993 H03231 H59605 H01642 AA852876 AA113758 AA626915 AA746952 AI161014 AA099554 R69057

Table 76C:

Pkey:	Unique number corresponding to an Eos probeset
Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
Strand:	Indicates DNA strand from which exons were predicted.
Nt_position:	Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
402230	9966312	Minus	29782-29932
403047	3540153	Minus	59793-59966
404029	7671252	Plus	108716-111112
404049	3668074	Minus	75765-78155
404682	9797231	Minus	40977-41150
404875	9801324	Plus	96588-96732,97722-97831
404877	1519284	Plus	1095-2107
404977	3738341	Minus	43081-43229
405932	7767812	Minus	123525-123713
406400	9256298	Plus	1553-1712,1878-2140,4252-4385,5922-6077

Table 77 provides Pkey; Seq ID No; Disease Indications; and Preferred Utility for sequences in Table 78. Seq ID No links the information in Table 77 to Table 78.

Seq ID No:	Sequence ID No for sequences in table	
Pkey:	Unique Eos probeset identifier number	
Disease Indications:	Diseases designated for coverage as described in Table 1	
Preferred Utility:	Preference of utility, based partly upon predicted localization (Ab is antibody; sm is small molecule target; CTL is vaccine target)	

Seq ID No	Pkey	Disease Indications	Preferred Utility
Seq ID 1 & 2	425023	angiogenesis	Ab, sm, CTL, imaging
Seq ID 3 & 4	424503	pancreas, prostate, angiogenesis, bladder, lung	Ab, sm, imaging
Seq ID 5 & 6	429423	angiogenesis	Ab, sm
Seq ID 7 & 8	400289	angiogenesis, bladder, lung, cervical, ovarian, head & neck	Ab, sm, CTL, diagnostic
Seq ID 9 & 10	419172	angiogenesis, renal	Ab, sm, CTL, imaging
Seq ID 11 & 12	418007	angiogenesis, lung, bladder, fibrosis, head & neck, pancreas, stomach, colon, ovarian	Ab, sm, diagnostic
Seq ID 13 & 14	418007	angiogenesis, lung, bladder, fibrosis, head & neck, pancreas, stomach, colon, ovarian	Ab, sm, diagnostic
Seq ID 15 & 16	407836	angiogenesis	CTL
Seq ID 17 & 18	414577	angiogenesis	Ab, CTL, diagnostic
Seq ID 19 & 20	416738	angiogenesis, lung, ovarian, bladder & stomach, pancreas, uterine	Ab, sm, CTL, imaging
Seq ID 21 & 22	428368	angiogenesis, head & neck, stomach	Ab, sm, diagnostic
Seq ID 23 & 24	415138	angiogenesis, pancreas, stomach, lung, uterine	Ab, CTL, diagnostic
Seq ID 25 & 26	429276	angiogenesis, bladder, glioblastoma	Ab, sm, imaging
Seq ID 27 & 28	416994	prostate, angiogenesis	Ab, CTL, imaging
Seq ID 29 & 30	407975	angiogenesis, renal	Ab, CTL, diagnostic
Seq ID 31 & 32	429113	angiogenesis, bladder and stomach	sm, CTL
Seq ID 33 & 34	418506	angiogenesis, ovarian, glioblastoma, uterine, lung, bladder, pancreas	Ab, sm, imaging
Seq ID 35 & 36	423961	breast, colon, bladder, lung, fibrosis, pancreas, head and neck, ovarian	Ab, sm, diagnostic
Seq ID 37 & 38	414812	breast, bladder, lung, fibrosis, pancreas, colon, head and neck, cervical, stomach, renal, ovarian	Ab, CTL, diagnostic
Seq ID 39 & 40	417433	pancreas, breast, bladder, lung, head & neck, cervical, uterine, ovarian, stomach, renal	Ab, CTL, imaging
Seq ID 41 & 42	417433	pancreas, breast, bladder, lung, head & neck, cervical, uterine, ovarian, stomach, renal	Ab, CTL, imaging
Seq ID 43 & 44	424399	breast, uterine, head & neck	Ab, CTL, diagnostic
Seq ID 45 & 46	422867	breast, ovarian, prostate, pancreas, lung, colon, uterine	Ab, CTL, diagnostic
Seq ID 47 & 48	428227	breast, lung, bladder, ovarian, head & neck, fibrosis, colon, stomach, cervical	Ab, CTL, diagnostic
Seq ID 49 & 50	444381	breast, colon, bladder, lung, pancreas, head & neck, ovarian, stomach, uterine, renal, angiogenesis	Ab, CTL, diagnostic
Seq ID 51 & 52	439569	breast, androgen withdrawal prostate, prostate, bladder	Ab, CTL, diagnostic
Seq ID 53 & 54	411558	pancreas, prostate, stomach, breast, uterine, cervical, ovarian	Ab, sm, imaging
Seq ID 55 & 56	400303	breast, ovarian, pros, stomach, uterine, bladder lung head & neck	Ab, sm, CTL, imaging
Seq ID 57 & 58	411789	pancreas, lung, breast, stomach, head & neck, ovarian, uterine	Ab, CTL, diagnostic
Seq ID 59 & 60	428698	breast, colon, lung, pancreas, stomach, head & neck, ovarian	Ab, sm, imaging
Seq ID 61 & 62	450098	breast, lung, stomach, uterine	CTL
Seq ID 63 & 64	421552	breast, ovarian, pancreas, cervical, uterine, prostate, lung, stomach, head & neck	Ab, sm, CTL, diagnostic
Seq ID 65 & 66	452747	breast, bladder, lung, head & neck, ovarian, stomach, uterine, pancreas	Ab, imaging
Seq ID 67 & 68	415539	breast, prostate, ovarian, cervical, uterine	Ab, sm, CTL, imaging
Seq ID 69 & 70	416636	breast, pancreas, uterine	Ab, sm, CTL, imaging
Seq ID 71 & 72	416636	breast, pancreas, uterine	Ab, sm, CTL, imaging
Seq ID 73 & 74	409079	breast, prostate	Ab, sm, CTL, imaging
Seq ID 75 & 76	442082	breast, prostate, ovarian	Ab, imaging
Seq ID 77 & 78	400297	breast, bladder, colon, prostate	Ab, sm, CTL, imaging
Seq ID 79 & 80	451398	breast, ovarian	CTL
Seq ID 81 & 82	429220	breast, prostate, benign prostatic hyperplasia	Ab, CTL, imaging
Seq ID 83 & 84	421524	breast	Ab, sm, CTL, imaging
Seq ID 85 & 86	423242	breast, renal, ovarian, prostate, colon	Ab, CTL, imaging
Seq ID 87 & 88	423242	breast, renal, ovarian, prostate, colon	CTL
Seq ID 89 & 90	423242	breast, renal, ovarian, prostate, colon	CTL
Seq ID 91 & 92	452190	breast, stomach, pancreas	CTL
Seq ID 93 & 94	452190	breast, stomach, pancreas	CTL
Seq ID 95 & 96	325372	breast	CTL
Seq ID 97 & 98	450375	breast, ovarian, head & neck, pancreas, lung, colon	Ab, sm, CTL, imaging
Seq ID 99 & 100	426215	breast, lung, renal, colon, ovarian, uterine	Ab, CTL, diagnostic
Seq ID 101 & 102	425247	breast, ovarian, lung, colon, pancreas, head & neck, stomach, uterine, cervical, bladder	Ab, sm, CTL, diagnostic
Seq ID 103 & 104	429353	breast, prostate	Ab, sm, CTL, imaging
Seq ID 105 & 106	429353	breast, prostate	Ab, sm, CTL, imaging
Seq ID 107 & 108	432201	breast, colon, lung chemo, ovarian, stomach, pancreas, uterine, cervical	Ab, sm, CTL, imaging
Seq ID 109 & 110	427585	breast, lung, head & neck, pancreas, stomach, colon, ovarian, cervical	CTL
Seq ID 111 & 112	446163	breast, cervical, uterine	Ab, sm, imaging
Seq ID 113 & 114	442117	breast, lung, bladder, pancreas, head & neck, stomach, ovarian, prostate	Ab, CTL, imaging
Seq ID 115 & 116	426179	breast	sm, CTL
Seq ID 117 & 118	431211	colon, bladder, lung, pancreas, head & neck	Ab, sm, imaging
Seq ID 119 & 120	447033	colon, prostate, fibrosis, breast, lung pancreas	Ab, CTL, diagnostic
Seq ID 121 & 122	447033	colon, prostate, fibrosis, breast, lung pancreas	Ab, CTL, diagnostic
Seq ID 123 & 124	447033	colon, prostate, fibrosis, breast, lung pancreas	Ab, CTL, diagnostic
Seq ID 125 & 126	115522	colon, lung, bladder, pancreas	CTL
Seq ID 127 & 128	452679	prostate, colon, pancreas, Taxol prostate	Ab, CTL, imaging
Seq ID 129 & 130	446051	colon, breast	Ab, sm, CTL, imaging
Seq ID 131 & 132	422048	colon, pancreas, prostate	diagnostic
Seq ID 133 & 134	410418	colon, bladder, lung, ovarian, pancreas, head & neck	Ab, sm, CTL, imaging
Seq ID 135 & 136	446342	uterine, colon, prostate	Ab, sm, CTL, imaging
Seq ID 137 & 138	422260	colon, ovarian mucinous	Ab, sm, CTL, diagnostic
Seq ID 139 & 140	409041	uterine, ovarian, lung, colon, stomach, head & neck, breast, pancreas	Ab, CTL, diagnostic
Seq ID 141 & 142	409041	uterine, ovarian, lung, colon, stomach, head & neck, breast, pancreas	Ab, CTL, diagnostic
Seq ID 143 & 144	111929	colon, breast, fibrosis	Ab, sm, imaging
Seq ID 145 & 146	111929	colon, breast, fibrosis	Ab, sm, imaging
Seq ID 147 & 148	111929	colon, breast, fibrosis	Ab, sm, imaging

Seq ID	Number	Tissue	Classification
Seq ID 149 & 150	104888	colon, stomach, uterine	Ab, imaging
Seq ID 151 & 152	420159	bladder, stomach	Ab, sm, CTL, imaging
Seq ID 153 & 154	422330	pancreas, colon, bladder	Ab,sm, CTL, imaging, diagnostic
Seq ID 155 & 156	452461	bladder, lung, head & neck, ovarian, glioblastoma, stomach, colon, cervical	CTL
Seq ID 157 & 158	413324	bladder	Ab, sm, CTL, diagnostic
Seq ID 159 & 160	412420	bladder, glioblastoma, lung, stomach	Ab, diagnostic
Seq ID 161 & 162	416658	lung, ovarian, uterine, bladder	Ab, CTL, diagnostic
Seq ID 163 & 164	407811	bladder, pancreas, stomach, uterine, lung	Ab, sm, diagnostic
Seq ID 165	402230	bladder	sm, CTL
Seq ID 166 & 167	402230	bladder	sm, CTL
Seq ID 168 & 169	432829	bladder	CTL
Seq ID 170 & 171	425721	bladder	Ab, imaging
Seq ID 172 & 173	420370	bladder	Ab, CTL, imaging
Seq ID 174 & 175	437852	bladder, lung	Ab, sm, CTL, imaging
Seq ID 176 & 177	402075	bladder, lung, head & neck, cervical	diagnostic
Seq ID 178 & 179	421110	bladder, pancreas, stomach, ovarian, lung	Ab, sm, diagnostic
Seq ID 180 & 181	451668	bladder, ovarian, lung	Ab, sm, CTL, diagnostic
Seq ID 182 & 183	451668	bladder, ovarian, lung	Ab, sm, CTL, diagnostic
Seq ID 184 & 185	451668	bladder, ovarian, lung	Ab, sm, CTL, diagnostic
Seq ID 186 & 187	408243	bladder, stomach, head & neck, cervical	Ab, diagnostic
Seq ID 188 & 189	422282	bladder, lung, head & neck	CTL, diagnostic
Seq ID 190 & 191	425852	bladder, lung, head & neck	Ab, sm, CTL, imaging
Seq ID 192 & 193	439738	bladder, lung, cervical	Ab, sm, CTL, imaging
Seq ID 194 & 195	404975	bladder	sm, CTL
Seq ID 196 & 197	425883	bladder, pancreas	Ab,CTL, imaging
Seq ID 198 & 199	404977	bladder, ovarian	Ab, sm, CTL, diagnostic
Seq ID 200 & 201	420876	pancreas, bladder	Ab,sm, CTL, imaging
Seq ID 202 & 203	427747	bladder, lung, ovarian, stomach	sm, CTL
Seq ID 204 & 205	420291	lung, bladder, ovarian, pancreas	Ab, sm, imaging
Seq ID 206 & 207	446673	bladder	sm, CTL
Seq ID 208	437553	bladder	Ab,CTL, imaging
Seq ID 209 & 210	437553	bladder	Ab,CTL, imaging
Seq ID 211 & 212	437553	bladder	Ab,CTL, imaging
Seq ID 213 & 214	426900	bladder, prostate	Ab,CTL, imaging
Seq ID 215 & 216	426900	bladder, prostate	Ab,CTL, imaging
Seq ID 217 & 218	426900	bladder, prostate	Ab,CTL, imaging
Seq ID 219 & 220	405932	bladder, lung, head & neck, cervical	sm
Seq ID 221 & 222	405932	bladder, lung, head & neck, cervical	sm
Seq ID 223 & 224	424008	bladder, head & neck, stomach, cervical	Ab,sm, CTL, imaging
Seq ID 225 & 226	424008	bladder, head & neck, stomach, cervical	Ab,sm, CTL, imaging
Seq ID 227 & 228	424008	bladder, head & neck, stomach, cervical	Ab,sm, CTL, imaging
Seq ID 229 & 230	444342	hepatitis C, lung, fibrosis, bladder	Ab, CTL, diagnostic
Seq ID 231 & 232	421379	breast, pancreas, head & neck, lung, stomach, bladder, cervical, colon	Ab, CTL, diagnostic
Seq ID 233 & 234	417079	bladder, lung, head & neck, cervical	Ab, diagnostic
Seq ID 235 & 236	412986	glioblastoma	Ab,sm, CTL, imaging
Seq ID 237 & 238	412986	glioblastoma	Ab,sm, CTL, imaging
Seq ID 239 & 240	412986	glioblastoma	Ab,sm, CTL, imaging
Seq ID 241 & 242	447072	glioblastoma, pancreas	sm, CTL
Seq ID 243 & 244	419723	glioblastoma	Ab, CTL, diagnostic
Seq ID 245 & 246	419723	glioblastoma	Ab,sm, CTL, imaging
Seq ID 247 & 248	430890	glioblastoma, lung, cervical, bladder	Ab, CTL, imaging, diagnostic
Seq ID 249 & 250	456759	glioblastoma	Ab,sm, CTL, imaging
Seq ID 251	429466	glioblastoma, uterine	
Seq ID 252	429466	glioblastoma, uterine	
Seq ID 253 & 254	419721	glioblastoma, lung	Ab,sm, CTL, imaging
Seq ID 255 & 256	407034	glioblastoma	Ab, CTL, diagnostic
Seq ID 257 & 258	413472	glioblastoma	Ab,sm, CTL, imaging
Seq ID 259 & 260	438380	glioblastoma	Ab, CTL, diagnostic, imaging
Seq ID 261 & 262	426271	glioblastoma	Ab, CTL, diagnostic
Seq ID 263 & 264	419704	glioblastoma	sm, CTL
Seq ID 265 & 266	444471	glioblastoma, lung, colon	Ab,sm, CTL, imaging
Seq ID 267 & 268	409395	glioblastoma	Ab, CTL, diagnostic
Seq ID 269 & 270	413063	glioblastoma, ovarian, bladder, lung	Ab, CTL, diagnostic
Seq ID 271 & 272	433800	glioblastoma, lung	Ab, CTL, imaging
Seq ID 273	458435	glioblastoma	
Seq ID 274	458435	glioblastoma	
Seq ID 275 & 276	424343	glioblastoma, ovarian, uterine	Ab,sm, CTL, imaging
Seq ID 277 & 278	424998	glioblastoma	Ab, CTL, diagnostic
Seq ID 279 & 280	412709	glioblastoma	Ab,sm, CTL, imaging
Seq ID 281 & 282	435615	glioblastoma	Ab,sm, CTL, imaging
Seq ID 283 & 284	404049	glioblastoma	Ab,sm, CTL, imaging
Seq ID 285 & 286	418932	glioblastoma	Ab,sm, CTL, imaging
Seq ID 287 & 288	404029	glioblastoma	Ab,sm, CTL, imaging
Seq ID 289 & 290	436480	glioblastoma	sm, CTL
Seq ID 291 & 292	452401	bladder, breast, pancreas, head & neck, stomach, lung, arthritis, renal	Ab, CTL, diagnostic
Seq ID 293 & 294	452401	bladder, breast, pancreas, head & neck, stomach, lung, arthritis, renal	Ab, CTL, diagnostic
Seq ID 295 & 296	436895	breast, renal, ovarian, glioblastoma	Ab,sm, imaging
Seq ID 297 & 298	421471	renal	Ab,sm, CTL, imaging
Seq ID 299 & 300	428296	renal	Ab,sm, CTL, imaging
Seq ID 301 & 302	423508	renal, colon	Ab, CTL, imaging
Seq ID 303 & 304	450001	renal, lung	Ab,sm, CTL, imaging
Seq ID 305 & 306	410407	renal, lung, colon, stomach, ovarian, uterine	Ab,sm, CTL, imaging
Seq ID 307 & 308	453496	renal, prostate	Ab,sm, CTL, imaging

Seq ID 309 & 310 420737	renal	Ab,sm, CTL, Imaging
Seq ID 311 & 312 309931	lung	Ab,sm, CTL, Imaging
Seq ID 313 & 314 412719	lung, head & neck, bladder, globlastoma, cervical	CTL
Seq ID 315 & 316 417034	lung, head & neck, a couple cervical	Ab, CTL, diagnostic
Seq ID 317 & 318 430486	lung, bladder, head & neck, cervical	Ab,sm, imaging
Seq ID 319 & 320 413753	lung, bladder, head & neck, pancreas, cervical, colon	CTL, diagnostic
Seq ID 321 & 322 425650	lung, head & neck, cervical, bladder	Ab, imaging
Seq ID 323 & 324 423673	bladder, lung, head & neck, ovarian, pancreas, colon, stomach, uterine, cervical	Ab, CTL, diagnostic
Seq ID 325 & 326 418663	lung, bladder, head & neck, cervical	Ab, imaging
Seq ID 327 & 328 418663	lung, bladder, head & neck, cervical	Ab, imaging
Seq ID 329 & 330 429610	lung	CTL, diagnostic
Seq ID 331 & 332 406690	lung, head & neck, pancreas, stomach, bladder, colon, cervical	Ab, imaging
Seq ID 333 & 334 431846	lung, bladder, head & neck, uterine, cervical, stomach, ovarian	Ab, imaging
Seq ID 335 & 336 422158	head & neck, bladder, lung, cervical, stomach	diagnostic
Seq ID 337 & 338 431958	lung, bladder, cervical, head & neck, ovarian, colon, prostate, pancreas, breast	Ab, CTL, imaging
Seq ID 339 & 340 437044	head & neck, cervical, lung, bladder, breast, prostate, ovarian, stomach	Ab, imaging
Seq ID 341 & 342 428484	lung, globlastoma, bladder, head & neck, colon, cervical	Ab, sm, imaging
Seq ID 343 & 344 429211	lung, bladder, head & neck, cervical, stomach	Ab, imaging
Seq ID 345 & 346 417389	ovarian, lung, bladder, uterine, cervical, pancreas, stomach	Ab, diagnostic
Seq ID 347 & 348 431009	lung, bladder, head & neck	Ab, sm, imaging
Seq ID 349 & 350 417542	lung	CTL, diagnostic
Seq ID 351 & 352 449230	lung, cervical, head & neck, bladder, ovarian, colon	Ab, imaging
Seq ID 353 & 354 410555	lung	Ab, sm, imaging
Seq ID 355 & 356 410555	lung	Ab, sm, imaging
Seq ID 357 & 358 424687	head & neck, pancreas, lung, uterine, cervical, colon, stomach	Ab, sm, diagnostic
Seq ID 359 & 360 418462	lung, bladder	Ab, imaging
Seq ID 361 & 362 410274	lung, renal	diagnostic
Seq ID 363 & 364 439606	lung, bladder, head & neck, cervical .	Ab,sm, imaging
Seq ID 365 & 366 404877	lung, bladder	CTL
Seq ID 367 & 368 444781	lung, bladder, head & neck, cervical	Ab, imaging
Seq ID 369 & 370 109424	bladder, lung	Ab, sm, CTL, imaging
Seq ID 371 & 372 415817	lung, globlastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
Seq ID 373 & 374 415817	lung, globlastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
Seq ID 375 & 376 415817	lung, globlastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
Seq ID 377 & 378 415817	lung, globlastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
Seq ID 379 & 380 415817	lung, globlastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
Seq ID 381 & 382 415817	lung, globlastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
Seq ID 383 & 384 421817	lung, cervical, head & neck & bladder	Ab,sm, CTL, imaging
Seq ID 385 & 386 418678	lung, bladder, stomach, ovarian, pancreas & cervical	CTL
Seq ID 387 & 388 418678	lung, bladder, stomach, ovarian, pancreas & cervical	CTL
Seq ID 389 & 390 409420	lung, head & neck, pancreas, stomach, cervical, bladder	CTL, diagnostic
Seq ID 391 & 392 332160	lung	Ab, sm, imaging
Seq ID 393 & 394 408790	lung	Ab,sm, CTL, imaging
Seq ID 395 & 396 408790	lung	Ab,sm, CTL, imaging
Seq ID 397 & 398 439223	lung, head & neck, cervical, bladder & colon	Ab, CTL, imaging
Seq ID 399 & 400 409757	pancreas, stomach, lung, bladder, stomach	Ab, CTL, diagnostic
Seq ID 401 & 402 428969	lung, cervical	Ab, CTL, diagnostic
Seq ID 403 & 404 428969	lung, cervical	Ab, CTL, diagnostic
Seq ID 405 & 406 428969	lung, cervical	Ab, CTL, diagnostic
Seq ID 407 & 408 428969	lung, cervical	Ab, CTL, diagnostic
Seq ID 409 & 410 450701	lung, head & neck, pancreas, breast, stomach, ovarian, colon	Ab, CTL, diagnostic
Seq ID 411 & 412 450701	lung, head & neck, pancreas, breast, stomach, ovarian, colon	Ab, CTL, diagnostic
Seq ID 413 & 414 414774	lung, bladder, head & neck, pancreas, stomach, ovarian	Ab, sm, diagnostic
Seq ID 415 & 416 407944	lung, head & neck	Ab, sm, imaging
Seq ID 417 & 418 407944	lung, head & neck	Ab, sm, imaging
Seq ID 419 & 420 428486	pancreas	Ab, CTL, diagnostic
Seq ID 421 & 422 457489	pancreas, prostate, lung	Ab, CTL, diagnostic
Seq ID 423 & 424 432874	pancreas, stomach	Ab, CTL, diagnostic
Seq ID 425 & 426 445891	stomach, pancreas, ovarian	Ab, CTL, imaging
Seq ID 427 & 428 445891	stomach, pancreas, ovarian	Ab, CTL, imaging
Seq ID 429 & 430 404682	pancreas	Ab, CTL, diagnostic
Seq ID 431 & 432 429547	pancreas, head & neck, lung, ovarian	Ab, CTL, diagnostic
Seq ID 433 & 434 425921	stomach, pancreas	Ab,sm, CTL, imaging
Seq ID 435 & 436 407242	pancreas, colon, bladder, head & neck, stomach, lung, ovarian, cervical	Ab,sm, CTL, imaging
Seq ID 437 & 438 407242	pancreas, colon, bladder, head & neck, stomach, lung, ovarian, cervical	Ab,sm, CTL, imaging
Seq ID 439 & 440 407242	pancreas, colon, bladder, head & neck, stomach, lung, ovarian, cervical	Ab,sm, CTL, imaging
Seq ID 441 & 442 432596	pancreas, breast	CTL
Seq ID 443 & 444 444006	pancreas, colon, lung, ovarian & cervical	Ab, CTL, imaging
Seq ID 445 & 446 423685	pancreas, uterine, colon	CTL
Seq ID 447 & 448 428392	pancreas	Ab, CTL, diagnostic
Seq ID 449 & 450 429597	pancreas, colon, stomach, lung	Ab,sm, CTL, imaging
Seq ID 451 & 452 448030	pancreas, renal and stomach,	Ab,sm, imaging
Seq ID 453 & 454 422109	pancreas, lung, colon	Ab, CTL, diagnostic
Seq ID 455 & 456 419235	pancreas, fibrosis, head & neck & lung	Ab, CTL, imaging
Seq ID 457 & 458 449048	pancreas, ovarian, uterine, globlastoma, head & neck & lung	Ab, CTL, imaging
Seq ID 459 & 460 427333	pancreas, colon	Ab, sm, imaging
Seq ID 461 & 462 417931	ovarian, pancreas, stomach, colon, uterine, prostate	Ab, diagnostic
Seq ID 463 & 464 419216	pancreas, lung, stomach, cervical, prostate, head & neck	Ab, CTL, diagnostic
Seq ID 465 & 466 431629	pancreas, uterine, cervical, stomach	Ab, CTL, diagnostic
Seq ID 467 & 468 413554	pancreas, globlastoma	Ab, CTL, diagnostic
Seq ID 469 & 470 452194	stomach, pancreas, renal, colon	Ab,sm, CTL, imaging
Seq ID 471 & 472 452194	stomach, pancreas, renal, colon	Ab,sm, CTL, imaging
Seq ID 473 & 474 426322	pancreas, bladder, stomach	Ab, diagnostic

Seq ID 475 & 476 429010	pancreas	sm, CTL
Seq ID 477 & 478 431462	pancreas, lung, glioblastoma	Ab, diagnostic
Seq ID 479 & 480 448243	ovarian, uterine, lung, stomach, head & neck, glioblastoma, pancreas	Ab,sm, imaging
Seq ID 481 & 482 428427	ovarian, lung, head & neck, cervical, colon, uterine, stomach	sm, CTL
Seq ID 483 & 484 428187	ovarian, uterine, colon, stomach	Ab,sm, CTL, imaging
Seq ID 485 & 486 428579	ovarian, Ewing, uterine, pancreas	Ab,sm, CTL, imaging
Seq ID 487 & 488 428579	ovarian, Ewing, uterine, pancreas	Ab,sm, CTL, imaging
Seq ID 489 & 490 428579	ovarian, Ewing, uterine, pancreas	Ab,sm, CTL, imaging
Seq ID 491 & 492 428579	ovarian, Ewing, uterine, pancreas	Ab,sm, CTL, imaging
Seq ID 493 & 494 445537	ovarian, bladder, uterine, breast, lung, head & neck, renal, fibrosis, pancreas, cervical	Ab, CTL, diagnostic
Seq ID 495 & 496 422278	ovarian, head & neck, bladder, cervical, lung	Ab,sm, imaging
Seq ID 497 & 498 424620	ovarian	Ab, CTL, diagnostic
Seq ID 499 & 500 406400	ovarian, uterine	Ab, CTL, diagnostic
Seq ID 501 & 502 431130	ovarian, uterine	Ab, CTL, diagnostic
Seq ID 503 & 504 420440	ovarian, uterine, cervical	Ab, CTL, diagnostic
Seq ID 505 & 506 428450	ovarian, cervical, pancreas, lung	sm
Seq ID 507 & 508 446619	ovarian, fibrosis, pancreas, head & neck, lung, colon	Ab, diagnostic
Seq ID 509 & 510 436982	ovarian, fibrosis	Ab, CTL, diagnostic
Seq ID 511 & 512 453392	ovarian, lung, glioblastoma	CTL
Seq ID 513 431969	ovarian	
Seq ID 514 439820	ovarian, uterine, cervical, breast, prostate	
Seq ID 515 & 516 409178	ovarian, breast	Ab, CTL, diagnostic
Seq ID 517 & 518 426514	ovarian, colon, bladder, lung, cervical	Ab, CTL, diagnostic
Seq ID 519 & 520 421478	ovarian, bladder	Ab, CTL, imaging
Seq ID 521 & 522 421478	ovarian, bladder	Ab, CTL, imaging
Seq ID 523 & 524 425776	ovarian, uterine, lung	Ab,sm, CTL, imaging
Seq ID 525 & 526 425776	ovarian, uterine, lung	Ab,sm, CTL, imaging
Seq ID 527 & 528 452097	ovarian	Ab, sm, diagnostic
Seq ID 529 & 530 416530	ovarian, uterine	Ab, CTL, diagnostic
Seq ID 531 & 532 431515	ovarian, prostate, lung	Ab,sm, CTL, imaging
Seq ID 533 & 534 419452	ovarian, prostate, lung, breast, uterine	Ab,sm, CTL, imaging
Seq ID 535 & 536 412078	ovarian	CTL
Seq ID 537 & 538 412078	ovarian	CTL
Seq ID 539 & 540 411773	ovarian	Ab, sm, CTL, diagnostic
Seq ID 541 & 542 407792	ovarian, uterine, cervical, pancreas	Ab, CTL, diagnostic
Seq ID 543 & 544 431616	prostate, pancreas, colon	Ab,sm, CTL, imaging
Seq ID 545 & 546 452792	prostate, uterine, breast	Ab, CTL, imaging
Seq ID 547 & 548 400294	prostate, taxol prostate	Ab,sm, CTL, imaging
Seq ID 549 & 550 432653	prostate, lung	Ab,sm, CTL, imaging
Seq ID 551 & 552 432653	prostate, lung	Ab,sm, CTL, imaging
Seq ID 553 & 554 432653	prostate, lung	Ab,sm, CTL, imaging
Seq ID 555 & 556 432653	prostate, lung	Ab,sm, CTL, imaging
Seq ID 557 & 558 400290	prostate, colon	Ab,sm, CTL, imaging
Seq ID 559 & 560 410001	ovarian, prostate, uterine, cervical, lung	Ab, CTL, diagnostic
Seq ID 561 & 562 418396	prostate	Ab,sm, CTL, imaging
Seq ID 563 & 564 451027	prostate, uterine, glioblastoma	Ab,sm, CTL, imaging
Seq ID 565 & 566 446057	prostate	Ab,sm, CTL, imaging
Seq ID 567 & 568 433466	prostate	Ab, CTL, diagnostic
Seq ID 569 & 570 453370	prostate	Ab,sm, CTL, imaging
Seq ID 571 & 572 453370	prostate	Ab,sm, CTL, imaging
Seq ID 573 & 574 414569	prostate	Ab,sm, CTL, imaging
Seq ID 575 & 576 413435	prostate, glioblastoma, pancreas	Ab, sm, diagnostic
Seq ID 577 & 578 426501	prostate, breast, glioblastoma, lung	Ab, CTL, imaging
Seq ID 579 & 580 448999	prostate, glioblastoma	Ab,sm, CTL, imaging
Seq ID 581 & 582 408369	prostate, lung, fibrosis, uterine, glioblastoma, cervical, ovarian	Ab,sm, CTL, imaging
Seq ID 583 & 584 412628	prostate	Ab, CTL, diagnostic
Seq ID 585 & 586 403047	prostate, bladder, colon	Ab,sm, CTL, imaging
Seq ID 587 & 588 403047	prostate, bladder, colon	Ab,sm, CTL, imaging
Seq ID 589 & 590 408430	prostate	Ab,sm, CTL, imaging
Seq ID 591 & 592 445413	prostate, colon, uterine, ovarian, lung, pancreas	diagnostic
Seq ID 593 & 594 451982	prostate, bladder	CTL
Seq ID 595 & 596 427953	prostate, glioblastoma	Ab,sm, CTL, imaging
Seq ID 597 & 598 421887	prostate	Ab, CTL, imaging
Seq ID 599 & 600 425071	prostate, colon, stomach, uterine, cervical, head & neck, pancreas	Ab, diagnostic
Seq ID 601 & 602 432101	prostate, pancreas	Ab,sm, imaging
Seq ID 603 & 604 407786	prostate, colon, uterine, stomach, inflammatory bowel disease, ovarian	Ab,sm, imaging
Seq ID 605 & 606 416836	Prostate, Ewing, glioblastoma	Ab, CTL, diagnostic
Seq ID 607 & 608 416539	ZD1839 resistant cancers, head & neck	Ab, sm, CTL, imaging
Seq ID 609 & 610 131083	androgen withdrawal prostate	Ab, sm, CTL, imaging
Seq ID 611 & 612 131083	androgen withdrawal prostate	Ab, sm, CTL, imaging
Seq ID 613 & 614 131083	androgen withdrawal prostate	Ab, sm, CTL, imaging
Seq ID 615 & 616 131083	androgen withdrawal prostate	Ab, sm, CTL, imaging
Seq ID 617 & 618 422424	bladder, pancreas, prostate, angiogenesis, colon, stomach, lung	Ab, CTL, diagnostic
Seq ID 619 & 620 426970	stomach, pancreas, colon	Ab,sm, imaging
Seq ID 621 & 622 428330	uterine, ovarian, fibrosis, prostate, pancreas, lung, bladder, head & neck	Ab, sm, CTL, diagnostic
Seq ID 623 & 624 439018	uterine, stomach, prostate	Ab,sm, CTL, imaging
Seq ID 625 & 626 420610	uterine, ovarian endometrioid, lung	CTL
Seq ID 627 & 628 425723	ovarian endometrioid, uterine, colon	Ab, CTL, diagnostic
Seq ID 629 & 630 458662	uterine, ovarian	CTL
Seq ID 631 & 632 418281	uterine, ovarian	Ab, CTL, diagnostic
Seq ID 633 & 634 429903	lung	sm

Table 78

Seq ID NO: 1 DNA sequence
Nucleic Acid Accession #: NM_001400
Coding sequence: 251..1399

```
      1         11         21         31         41         51
      |          |          |          |          |          |
TCTAAAGGTC GGGGGCAGCA GCAAGATGCG AAGCGAGCCG TACAGATCCC GGGCTCTCCG    60
AACGCAACTT CGCCCTGCTT GAGCGAGGCT GCGGTTTCCG AGGCCCTCTC CAGCCAAGGA   120
AAAGCTACAC AAAAAGCCTG GATCACTCAT CGAACCACCC CTGAAGCCAG TGAAGGCTCT   180
CTCGCCTCGC CCTCTAGCGT TCGTCTGGAG TAGCGCCACC CCGGCTTCCT GGGGACACAG   240
GGTTGGCACC ATGGGGCCCA CCAGCGTCCC GCTGGTCAAG GCCCACCGCA GCTCGGTCTC   300
TGACTACGTC AACTATGATA TCATCGTCCG GCATTACAAC TACACGGGAA AGCTGAATAT   360
CAGCGCGGAC AAGGAGAACA GCATTAAACT GACCTCGGTG GTGTTCATTC TCATCTGCTG   420
CTTTATCATC CTGGAGAACA TCTTTGTCTT GCTGACCATT TGGAAAACCA AGAAATTCCA   480
CCGACCCATG TACTATTTTA TTGGCAATCT GGCCCTCTCA GACCTGTTGG CAGGAGTAGC   540
CTACACAGCT AACCTGCTCT TGTCTGGGGC CACCACCTAC AAGCTCACTC CCGCCCAGTG   600
GTTTCTGCGG GAAGGGAGTA TGTTTGTGGC CCTGTCAGCC TCCGTGTTCA GTCTCCTCGC   660
CATCGCCATT GAGCGCTATA TCACAATGCT GAAAATGAAA CTCCACAACG GGAGCAATAA   720
CTTCCGCCTC TTCCTGCTAA TCAGCGCCTG CTGGGTCATC TCCCTCATCC TGGGTGGCCT   780
GCCTATCATG GGCTGGAACT GCATCAGTGC GCTGTCCAGC TGCTCCACCG TGCTGCCGCT   840
CTACCACAAG CACTATATCC TCTTCTGCAC CACGGTCTTC ACTCTGCTTC TGCTCTCCAT   900
CGTCATTCTG TACTGCAGAA TCTACTCCTT GGTCAGGACT CGGAGCCGCC GCCTGACGTT   960
CCGCAAGAAC ATTTCCAAGG CCAGCCGCAG CTCTGAGAAG TCGCTGGCGC TGCTCAAGAC  1020
CGTAATTATC GTCCTGAGCG TCTTCATCGC CTGCTGGGCA CCGCTCTTCA TCCTGCTCCT  1080
GCTGGATGTG GGCTGCAAGG TGAAGACCTG TGACATCCTC TTCAGAGCGG AGTACTTCCT  1140
GGTGTTAGCT GTGCTCAACT CCGGCACCAA CCCCATCATT TACACTCTGA CCAACAAGGA  1200
GATGCGTCGG GCCTTCATCC GGATCATGTC CTGCTGCAAG CCGTCGGGGG ACTCTGC     1260
TGGCAAATTC AAGCGACCCA TCATCGCCGG CATGGAATTC AGCCGCAGCA AATCGGACAA  1320
TTCCTCCCAC CCCCAGAAAG ACGAAGGGGA CAACCCAGAG ACCATTATGT CTTCTGGAAA  1380
CGTCAACTCT TCTTCCTAGA ACTGGAAGCT GTCCACCCAC CGGAAGCGCT CTTTACTTGG  1440
TCGCTGGCCA CCCCAGTGTT TGGAAAAAAA TCTCTGGGCT TCGACTGCTG CCAGGGAGGA  1500
GCTGCTGCAA GCCAGAGGGA GGAAGGGGGA GAATACGAAC AGCCTGGTGG TGTCGGGTGT  1560
TGGTGGGTAG AGTTAGTTCC TGTGAACAAT GCACTGGGAA GGGTGGAGAT CAGGTCCCGG  1620
CCTGGAATAT ATATTCTACC CCCCTGGAGC TTTGATTTTG CACTGAGCCA AAGGTCTAGC  1680
ATTGTCAAGC TCCTAAGGGG TTCATTTGGC CCCTCCTCAA AGACTAATGT CCCCATGTGA  1740
AAGCGTCTCT TTGTCTGGAG CTTTGAGGAG ATGTTTTCCT TCACTTTAGT TTCAAACCCA  1800
AGTGAGTGTG TGCACTTCTG CTTCTTTAGG GATGCCCTGT ACATCCCACA CCCCACCCTC  1860
CCTTCCCTTC ATACCCCTCC TCAACGTTCT TTTACTTTAT ACTTTAACTA CCTGAGAGTT  1920
ATCAGAGCTG GGGTTGTGGA ATGATCGATC ATCTATAGCA AATAGGCTAT GTTGAGTACG  1980
TAGGCTGTGG GAAGATGAAG ATGGTTTGGA GGTGTAAAAC AATGTCCTTC GCTGAGGCCA  2040
AAGTTTCCAT GTAAGCGGGA TCCGTTTTTT GGAATTTGGT TGAAGTCACT TTGATTTCTT  2100
TAAAAAACAT CTTTTCAATG AAATGTGTTA CCATTTCATA TCCATTGAAG CCGAAATCTG  2160
CATAAGGAAG CCCACTTTAT CTAAATGATA TTAGCCAGGA TCCTTGGTGT CCTAGGAGAA  2220
ACAGACAAGC AAAACAAAGT GAAAACCGAA TGGATTAACT TTTGCAAACC AAGGGAGATT  2280
TCTTAGCAAA TGAGTCTAAC AAATATGACA TCCGTCTTTC CCACTTTTGT TGATGTTTAT  2340
TTCAGAATCT TGTGTGATTC ATTTCAAGCA ACAACATGTT GTATTTTGTT GTGTTAAAAG  2400
TACTTTTCTT GATTTTTGAA TGTATTTGTT TCAGGAAGAA GTCATTTTAT GGATTTTTCT  2460
AACCCGTGTT AACTTTTCTA GAATCCACCC TCTTGTGCCC TTAAGCATTA CTTTAACTGG  2520
TAGGGAACGC CAGAACTTTT AAGTCCAGCT ATTCATTAGA TAGTAATTGA AGATATGTAT  2580
AAAATATTACA AAGAATAAAA ATATATTACT GTCTCTTTAG TATGGTTTTC AGTGCAATTA  2640
AACCGAGAGA TGTCTTGTTT TTTTAAAAAG AATAGTATTT AATAGGTTTC TGACTTTTGT  2700
GGATCATTTT GCACATAGCT TTATCAACTT TTAAACATTA ATAAACTGAT TTTTTTAAAG
```

Seq ID NO: 2 Protein sequence
Protein Accession #: NP_001391

```
      1         11         21         31         41         51
      |          |          |          |          |          |
MGPTSVPLVK AHRSSVSDYV NYDIIVRHYN YTGKLNISAD KENSIKLTSV VFILICCFII   60
LENIFVLLTI WKTKKFHRPM YYFIGNLALS DLLAGVAYTA NLLLSGATTY KLTPAQWFLR  120
EGSMFVALSA SVFSLLAIAI ERYITMLKMK LHNGSNNFRL FLLISACWVI SLILGGLPIM  180
GWNCISALSS CSTVLPLYHK HYILFCTTVF TLLLLSIVIL YCRIYSLVRT RSRRLTFRKN  240
ISKASRSSEK SLALLKTVII VLSVFIACWA PLFILLLLDV GCKVKTCDIL FRAEYFLVLA  300
VLNSGTNPII YTLTNKEMRR AFIRIMSCCK CPSGDSAGKF KRPIIAGMEF SRSKSDNSSH  360
PQKDEGDNPE TIMSSGNVNS SS
```

Seq ID NO: 3 DNA sequence
Nucleic Acid Accession #: NM_002205.1
Coding sequence: 1..3149

```
      1         11         21         31         41         51
      |          |          |          |          |          |
ATGGGGAGCC GGACGCCAGA GTCCCCTCTC CACGCCGTGC AGCTGCGCTG GGGCCCCCGG   60
CGCGACCCC CGCTSSTGCC GCTGCTGTTG CTGCTSSTGC CGCCGCCACC CAGGGTCGGG  120
GGCTTCAACT TAGACGCGGA GGCCCCAGCA GTACTCTCGG GGCCCCCGGG CTCCTCTTC  180
GGATTCTCAG TGGAGTTTTA CCGGCCGGGA ACAGACGGGG TCAGTGTGCT GGTGGGAGCA  240
CCCAAGGCTA ATACCAGCCA GCCAGGAGTG CTGCAGGGTG GTGCTGTCTA CCTCTGTCCT  300
TGGGGTGCCA GCCCCACACA GTGCACCCCC ATTGAATTTG ACAGCAAAGG CTCTCGGCTC  360
CTGGAGTCCT CACTGTCCAG CTCAGAGGGA GAGGAGCCTG TGGAGTACAA GTCCTTGCAG  420
TGGTTCGGGG CAACAGTTCG AGCCCATGGC TCCTCCATCT TGGCATGCGC TCCACTGTAC  480
AGCTGGCGCA CAGAGAAGGA GCCACTGAGC GACCCCGTGG GCACCTGCTA CCTCTCCACA  540
GATAACTTCA CCCGAATTCT GGAGTATGCA CCCTGCCGCT CAGATTTCAG CTGGGCAGCA  600
```

949

```
GGACAGGGTT  ACTGCCAAGG  AGGCTTCAGT  GCCGAGTTCA  CCAAGACTGG  CCGTGTGGTT   660
TTAGGTGGAC  CAGGAAGCTA  TTTCTGGCAA  GGCCAGATCC  TGTCTGCCAC  TCAGGAGCAG   720
ATTGCAGAAT  CTTATTACCC  CGAGTACCTG  ATCAACCTGG  TTCAGGGGCA  GCTGCAGACT   780
CGCCAGGCCA  GTTCCATCTA  TGATGACAGC  TACCTAGGAT  ACTCTGTGGC  TGTTGGTGAA   840
TTCAGTGGTG  ATGACACAGA  AGACTTTGTT  GCTGGTGTGC  CCAAAGGGAA  CCTCACTTAC   900
GGCTATGTCA  CCATCCTTAA  TGGCTCAGAC  ATTCGATCCC  TCTACAACTT  CTCAGGGGAA   960
CAGATGGCCT  CCTACTTTGG  CTATGCAGTG  GCCGCCACAG  ACGTCAATGG  GGACGGGCTG  1020
GATGACTTGC  TGGTGGGGGC  ACCCCTGCTC  ATGGATCGGA  CCCCTGACGG  GCGGCCTCAG  1080
GAGGTGGGCA  GGGTCTACGT  CTACCTGCAG  CACCCAGCCG  GCATAGAGCC  CACGCCCACC  1140
CTTACCCTCA  CTGGCCATGA  TGAGTTTGGC  CGATTTGGCA  GCTCCTTGAC  CCCCCTGGGG  1200
GACCTGGACC  AGGATGGCTA  CAATGATGTG  GCCATCGGGG  CTCCCTTTGG  TGGGGAGACC  1260
CAGCAGGGAG  TAGTGTTTGT  ATTTCCTGGG  GGCCCAGGAG  GGCTGGGCTC  TAAGCCTTCC  1320
CAGGTTCTGC  AGCCCCTGTG  GGCAGCCAGC  CACACCCCAG  ACTTCTTTGG  CTCTGCCCTT  1380
CGAGGAGGCC  GAGACCTGGA  TGGCAATGGA  TATCCTGATC  TGATTGTGGG  GTCCTTTGGT  1440
GTGGACAAGG  CTGTGGTATA  CAGGGGCCGC  CCCATCGTGT  CCGCTAGTGC  CTCCCTCACC  1500
ATCTTCCCCG  CCATGGTTCAA  CCCAGAGGAG  CGGAGCCTGCA  GCTTAGAGGG  GAACCCTGTG  1560
GCCTGCATCA  ACCTTAGCTT  CTGCCTCAAT  GCTTCTGGAA  AACACGTTGC  TGACTCCATT  1620
GGTTTCACAG  TGGAACTTCA  GCTGGACTGG  CAGAAGCAGA  AGGGAGGGGT  ACGGCGGGCA  1680
CTGTTCCTGG  CCTCCAGGCA  GGCAACCCTG  ACCCAGACCC  TGCTCATCCA  GAATGGGGCT  1740
CGAGAGGATT  GCAGAGAGAT  GAAGATCTAC  CTCAGGAACG  AGTCAGAATT  CGAGACAAA  1800
CTCTCGCCGA  TTCACATCGC  TCTCAACTTC  TCCTTGGACC  CCCAGCCCCC  AGTGGACAGC  1860
CACGGCCTCA  GGCCAGCCCT  ACATTATCAG  AGCAAGAGCC  GGATAGAGGA  CAAGGCTCAG  1920
ATCTTGCTGG  ACTGTGGAGA  AGACAACATC  TGTGTGCCTG  ACCTGCAGCT  GGAAGTGTTT  1980
GGGGGAGCAGA  ACCATGTGTA  CCTGGGTGAC  AAGAATGCCC  TGAACCTCAC  TTTCCATGCC  2040
CAGAATGTGG  GTGAGGGTGG  CGCCTATGAG  GCTGAGCTTC  GGGTCACCGC  CCCTCCAGAG  2100
GCTGAGTACT  CAGGACTCGT  CAGACACCCA  GGGAACTTCT  CCAGCCTGAG  CTGTGAACTAC  2160
TTTGCCGTGA  ACCAGAGCCG  CCTGCTGGTG  TGTGACCTGG  GCAACCCCAT  GAAGGCAGGA  2220
GCCAGTCTGT  GGGGTGGCCT  TCGGTTTACA  GTCCCTCATC  TCCGGGACAC  TAAGAAAACC  2280
ATCCAGTTTG  ACTTCCAGAT  CCTCAGCAAG  AATCTCAACA  ACTCGCAAAG  CGACGTGGTT  2340
TCCTTTCGGC  TCTCCGTGGA  GGCTCAGGCC  CAGGTCACCC  TGAACGGTGT  CTCCAAGCCT  2400
GAGGCAGTGC  TATTCCCAGT  AAGCGACTGG  CATCCCCGAG  ACCAGCCTCA  GAAGGAGGAG  2460
GACCTGGGAC  CTGCTGTCCA  CCATGTCTAT  GAGCTCATCA  ACCAAGGCCC  CAGCTCCATT  2520
AGCCAGGGTG  TGCTGGAACT  CAGCTGTCCC  CAGGCTCTGG  AAGGTCAGCA  GCTCCTATAT  2580
GTGACCAGAG  TTACGGGACT  CAACTGCACC  ACCAATCACC  CCATTAACCC  AAAGGGCCTG  2640
GAGTTGGATC  CCGAGGGGTTC  CCTGCACCAC  CAGCAAAAAC  GGGAAGCTCC  AAGCCGCAGC  2700
TCTGCTTCCT  CGGGACCTCA  GATCCTGAAA  TGCCCGGAGG  CTGAGTGTTT  CAGGCTGCGC  2760
TGTGAGCTCG  GGCCCCTGCA  CCAACAAGAG  AGCCAAAGTC  TGCAGTTGCA  TTTCCGAGTC  2820
TGGGCCAAGA  CTTTCTTGCA  GCGGGAGCAC  CAGCCATTTA  GCCTGCAGTG  TGAGGCTGTG  2880
TACAAAGCCC  TGAAGATGCC  CTACCGAATC  CTGCCTCGGC  AGCTGCCCCA  AAAAGAGCGT  2940
CAGGTGGCCA  CAGCTGTGCA  ATGGACCAAG  GCAGAAGGCA  GCTATGGCGT  CCCACTGTGG  3000
ATCATCATCC  TAGCCATCCT  GTTTGGCCTC  CTGCTCCTAG  GTCTACTCAT  CTACATCCTC  3060
TACAAGCTTG  GATTCTTCAA  ACGCTCCCTC  CCATATGGCA  CCGCCATGGA  AAAAGCTCAG  3120
CTCAAGCCTC  CAGCCACCTC  TGATGCCTGA
```

Seq ID NO: 4 <u>Protein sequence</u>
Protein Accession #: NP_002196.1

```
1           11          21          31          41          51
|           |           |           |           |           |
MGSRTPESPL  HAVQLRWGPR  RRPPLLPLLL  LLLPPPPRVG  GFNLDAEAPA  VLSGPPGSFF    60
GFSVEFYRPG  TDGVSVLVGA  PKANTSQPGV  LQGGAVYLCP  WGASPTQCTP  IEFDSKGSRL   120
LESSLSSSEG  EEPVEYKSLQ  WFGATVRAHG  SSILACAPLY  SWRTEKEPLS  DPVGTCYLST   180
DNFTRILEYA  PCRSDPSWAA  GQGYCQGGFS  AEFTKTGRVV  LGGPGSYFWQ  GQILSATQEQ   240
IAESYYPEYL  INLVQGQLQT  RQASSIYDDS  YLGYSVAVGE  FSGDDTEDFV  AGVPKGNLTY   300
GYVTILNGSD  IRSLYNFSGE  QMASYFGYAV  AATDVNGDGL  DDLLVGAPLL  MDRTPDGRPQ   360
EVGRVYVYLQ  HPAGIEPTPT  LTLTGHDEFG  RFGSSLTPLG  DLDQDGYNDV  AIGAPFGGET   420
QQGVVFVFPG  GPGGLGSKPS  QVLQPLWAAS  HTPDFFGSAL  RGGRDLDGNG  YPDLIVGSFG   480
VDKAVVYRGR  PIVSASASLT  IFPAMFNPEE  RSCSLEGNPV  ACINLSFCLN  ASGKHVADSI   540
GFTVELQLDW  QKQKGGVRRA  LFLASRQATL  TQTLLIQNGA  REDCREMKIY  LRNESEFRDK   600
LSPIHIALNF  SLDPQAPVDS  HGLRPALHYQ  SKSRIEDKAQ  ILLDCGEDNI  CVPDLQLEVF   660
GEQNHVYLGD  KNALNLTFHA  QNVGEGGAYE  AELRVTAPPE  AEYSGLVRHP  GNFSSLSCDY   720
FAVNQSRLLV  CDLGNPMKAG  ASLNGGLRFT  VPHLRDTKKT  IQFDFQILSK  NLNNSQSDVV   780
SFRLSVEAQA  QVTLNGVSKP  EAVLFPVSDW  HPRDQPQKEE  DLGPAVHHVY  ELINQGPSSI   840
SQGVLELSCP  QALEGQQLLY  VTRVTGLNCT  TNHPINPKGL  ELDPEGSLHH  QQKREAPSRS   900
SASSGPQILK  CPEAECFRLR  CELGPLHQQE  SQSLQLHFRV  WAKTFLQREH  QPFSLQCEAV   960
YKALKMPYRI  LPRQLPQKER  QVATAVQWTK  AEGSYGVPLW  IIILAILFGL  LLLGLLIYIL  1020
YKLGFFKRSL  PYGTAMEKAQ  LKPPATSDA
```

Seq ID NO: 5 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_002211.1
Coding sequence: 1..2397

```
1           11          21          31          41          51
|           |           |           |           |           |
ATGAATTTAC  AACCAATTTT  CTGGATTGGA  CTGATCAGTT  CAGTTTGCTG  TGTGTTTGCT    60
CAAACAGATG  AAAATAGATG  TTTAAAAGCA  AATGCCAAAT  CATGTGGAGA  ATGTATACAA   120
GCAGGGCCAA  ATGTGGGGTG  GTGCACAAAT  TCAACATTTT  TACAGGAAGG  AATGCCTACT   180
TCTGCACGAT  GTGATGATTT  AGAAGCCTTA  AAAAAGAAGG  GTTGCCCTCC  AGATGACATA   240
GAAATCCCA  GAGGCTCCAA  AGATATAAAG  AAAAATAAAA  ATGTAACCAA  CCGTAGCAAA   300
GGAACAGCGA  AGAAGCTCAA  GCCAGAGGAT  ATTACTCAGA  TCCACCACA  GCAGTTGGTT   360
TTGCGATTAA  GATCAGGGGA  GCCACAGACA  TTTACATTAA  AATTCAAGAG  AGCTGAAGAC   420
TATCCCATTG  ACCTCTACTA  CCTTATGGAC  CTGTCTTACT  CAATGAAAGA  CGATTTGGAG   480
AATGTAAAAA  GTCTTGGAAC  AGATCTGATG  AATGAAATGA  GGAGGATTAC  TTCGGACTTC   540
```

```
AGAATTGGAT TTGGCTCATT TGTGGAAAAG ACTGTGATGC CTTACATTAG CACAACACCA    600
GCTAAGCTCA GGAACCCTTG CACAAGTGAA CAGAACTGCA CCAGCCCATT TAGCTACAAA    660
AATGTGCTCA GTCTTACTAA TAAAGGAGAA GTATTTAATG AACTTGTTGG AAAACAGCGC    720
ATATCTGGAA ATTTGGATTC TCCAGAAGGT GGTTTCGATG CCATCATGCA AGTTGCAGTT    780
TGTGGATCAC TGATTGGCTG GAGGAATGTT ACACGGCTGC TGGTGTTTTC CACAGATGCC    840
GGGTTTCACT TTGCTGGAGA TGGGAAACTT GGTGGCATTG TTTTACCAAA TGATGGACAA    900
TGTCACCTGG AAAATAATAT GTACACAATG AGCCATTATT ATGATTATCC TTCTATTGCT    960
CACCTTGTCC AGAAACTGAG TGAAAATAAT ATTCAGACAA TTTTTGCAGT TACTGAAGAA    1020
TTTCAGCCTG TTTACAAGGA GCTGAAAAAC TTGATCCCTA AGTCAGCAGT AGGAACATTA    1080
TCTGCAAATT CTAGCAATGT AATTCAGTTG ATCATTGATG CATACAATTC CCTTTCCTCA    1140
GAAGTCATTT TGGAAAACGG CAAATTGTCA GAAGGAGTAA CAATAAGTTA CAAATCTTAC    1200
TGCAAGAACG GGGTGAATGG AACAGGGGAA AATGGAAGAA AATGTTCCAA TATTTCCATT    1260
GGAGATGAGG TTCAATTTGA AATTAGCATA ACTTCAAATA AGTGTCCAAA AAAGGATTCT    1320
GACAGCTTTA AAATTAGGCC TCTGGGCTTT ACGGAGGAAG TAGAGGTTAT TCTTCAGTAC    1380
ATCTGTGAAT GTGAATGCCA AAGCGAAGGC ATCCCTGAAA GTCCCAAGTG TCATGAAGGA    1440
AATGGGACAT TTGAGTGTGG CGCGTGCAGG TGCAATGAAG GGCGTGTTGG TAGACATTGT    1500
GAATGCAGCA CAGATGAAGT TAACAGTGAA GACATGGATG CTTACTGCAG GAAAGAAAAC    1560
AGTTCAGAAA TCTGCAGTAA CAATGGAGAG TGCGTCTGCG GACAGTGTGT TTGTAGGAAG    1620
AGGGATAATA CAAATGAAAT TTATTCTGGC AAATTCTGCG AGTGTGATAA TTTCAACTGT    1680
GATAGATCCA ATGGCTTAAT TTGTGGAGGA AATGGTGTTT GCAAGTGTCG TGTGTGTGAG    1740
TGCAACCCCA ACTACACTGG CAGTGCATGT GACTGTTCTT TGGATACTAG TACTTGTGAA    1800
GCCAGCAACG GACAGATCTG CAATGGCCGG GGCATCTGCG AGTGTGGTGT CTGTAAGTGT    1860
ACAGATCCGA AGTTTCAAGG GCAAACGTGT GAGATGTGTC AGACCTGCCT TGGTGTCTGT    1920
GCTGAGCATA AAGAATGTGT TCAGTGCAGA GCCTTCAATA AAGGAGAAAA GAAAGACACA    1980
TGCACACAGG AATGTTCCTA TTTTAACATT ACCAAGGTAG AAAGTCGGGA CAAATTACCC    2040
CAGCCGGTCC AACCTGATCC TGTGTCCCAT TGTAAGGAGA AGGATGTTGA CGACTGTTGG    2100
TTCTATTTTA CGTATTCAGT GAATGGGAAC AACGAGGTCA TGGTTCATGT TGTGGAGAAT    2160
CCAGAGTGTC CCACTGGTCC AGACATCATT CCAATTGTAG CTGGTGTGGT TGCTGGAATT    2220
GTTCTTATTG GCCTTGCATT ACTGCTGATA TGGAAGCTTT TAATGATAAT TCATGACAGA    2280
AGGGAGTTTG CTAAATTTGA AAAGGAGAAA ATGAATGCCA AATGGGACAC GGGTGAAAAT    2340
CCTATTTATA AGAGTGCCGT AACAACTGTG GTCAATCCGA AGTATGAGGG AAAATGA
```

Seq ID NO: 6 Protein sequence
Protein Accession #: NP_002202.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MNLQPIFWIG LISSVCCVFA QTDENRCLKA NAKSCGBCIQ AGPNCGWCTN STFLQBGMPT    60
SARCDDLEAL KKKGCPPDDI ENPRGSKDIK KNKNVTNRSK GTAEKLKPED ITQIQPQQLV    120
LRLRSGEPQT FTLKFKRAED YPIDLYYLMD LSYSMKDDLE NVKSLGTDLM NEMRRITSDF    180
RIGFGSFVEK TVMPYISTTP AKLRNPCTSE QNCTSPPSYK NVLSLTNKGE VFNELVGKQR    240
ISGNLDSPEG GFDAIMQVAV CGSLIGWRNV TRLLVFSTDA GFHFAGDGKL GGIVLPNDGQ    300
CHLENNMYTM SHYYDYPSIA HLVQKLSENN IQTIFAVTEE FQPVYKELKN LIPKSAVGTL    360
SANSSNVIQL IIDAYNSLSS EVILENGKLS EGVTISYKSY CKNGVNGTGE NGRKCSNISI    420
GDEVQFEISI TSNKCPKKDS DSPKIRPLGF TEEVEVILQY ICECECQSEG IPESPKCHEG    480
NGTFECGACR CNEGRVGRHC ECSTDEVNSE DMDAYCRKEN SSEICSNNGE CVCGQCVCRK    540
RDNTNEIYSG KFCECDNFNC DRSNGLICGG NGVCKCRVCE CNPNYTGSAC DCSLDTSTCS    600
ASNGQICNGR GICECGVCKC TDPKFQGQTC EMCQTCLGVC AEHKECVQCR AFNKGEKKDT    660
CTQECSYFNI TKVESRDKLP QPVQPDPVSH CKEKDVDDCW FYFTYSVNGN NEVMVHVVEN    720
PECPTGPDII PIVAGVVAGI VLIGLALLLI WKLLMIIHDR REFAKFEKEK MNAKWDTGEN    780
PIYKSAVTTV VNPKYEGK
```

Seq ID NO: 7 DNA sequence
Nucleic Acid Accession #: NM_002425
Coding sequence: 26..1453

```
1          11         21         31         41         51
|          |          |          |          |          |
AAAGAAGGTA AGGGCAGTGA GAATGATGCA TCTTGCATTC CTTGTGCTGT TGTGTCTGCC    60
AGTCTGCTCT GCCTATCCTC TGAGTGGGGC AGCAAAGAG GAGGACTCCA ACAAGGATCT    120
TGCCCAGCAA TACCTAGAAA AGTACTACAA CCTCGAAAAG GATGTGAAAC AGTTTAGAAG    180
AAAGGACAGT AATCTCATTG TTAAAAAAAT CCAAGGAATG CAGAAGTTCC TTGGGTTGGA    240
GGTGACAGGG AAGCTAGACA CTGACACTCT GGAGGTGATG CGCAAGCCCA GGTGTGGAGT    300
TCCTGACGTT GGTCACTTCA GCTCCTTTCC TGGCATGCCG AAGTGGAGGA AAACCCACCT    360
TACATACAGG ATTGTGAATT ATACACCAGA TTTGCCAAGA GATGCTGTTG ATTCTGCCAT    420
TGAGAAAGCT CTGAAAGTCT GGGAAGAGGT GACTCCACTC ACATTCTCCA GGCTGTATGA    480
AGGAGAGGCT GATATAATGA TCTCTTTCGC AGTTAAAGAA CATGGAGACT TTTACTCTTT    540
TGATGGCCCA GGACACAGTT TGGCTCATGC CTACCCACCT GGACCTGGGC TTTATGGAGA    600
TATTCACTTT GATGATGATG AAAAATGGAC AGAAGATGCA TCAGGCACCA ATTTATTCCT    660
CGTTGCTGCT CATGAACTTG GCCACTCCCT GGGGCTCTTT CACTCAGCCA ACACTGAAGC    720
TTTGATGTAC CCACTCTACA ACTCATTCAC AGAGCTCGCC CAGTTCCGCC TTTCGCAAGA    780
TGATGTGAAT GGCATTCAGT CTCTCTACGG ACCTCCCCGT GCCTCTACTG AGGAACCCGT    840
GGTGCCCACA AAATCTGTTC CTTCGGGATC TGAGATGCCA GCCAAGTGTG ATCCTCGCTTT    900
GTCCTTCGAT GCCATCAGCA CTCTGAGGGG AGAATATCTG TTCTTTAAAG ACAGATATTT    960
TTGGCGAAGA TCCCACTGGA ACCCTGAACC TGAATTTCAT TTGATTTCTG CATTTTGGCC    1020
CTCTCTTCCA TCATATTTGG ATGCTGCATA TGAAGTTAAC AGCAGGGACA CCGTTTTTAT    1080
TTTTAAGGGA AATGAGTTCT GGGCCATCAG AGGAAATGAG GTACAAGCAG GTTATCCAAG    1140
AGGCATCCAT ACCCTGGGTT TTCCTCCAAC CATAAGGAAA ATTGATGCAG CTGTTTCTGA    1200
CAAGGAAAAG AAGAAAACAT ACTTCTTTGC AGCGGACAAA TACTGGGAGA TTGATGAAAA    1260
TAGCCAGTCC ATGGAGCAAG GCTTCCCTAG ACTAATAGCT GATGACTTTC AGGAGTTGGA    1320
GCCTAAGGTT GATGCTGTAT TACAGGCATT TGGATTTTTC TACTTCTTCA GTGGATCATC    1380
ACAGTTTGAG TTTGACCCCA ATGCCAGGAT GGTGACACAC ATTAAAGA GTAACAGCTG    1440
GTTACATTGC TAGGCGAGAT AGGGGGAAGA CAGATATGGG TGTTTTTAAT AAATCTAATA    1500
ATTATTCATC TAATGTATTA TGAGCCAAAA TGGTTAATTT TTCCTGCATG TTCTGTGACT    1560
```

```
     GAAGAAGATG AGCCTTGCAG ATATCTGCAT GTGTCATGAA GAATGTTTCT GGAATTCTTC 1620
     ACTTGCTTTT GAATTGCACT GAACAGAATT AAGAAATACT CATGTGCAAT AGGTGAGAGA 1680
     ATGTATTTTC ATAGATGTGT TATTACTTCC TCAATAAAAA GTTTTATTTT GGGCCTGTTC 1740
5    CTT
```

Seq ID NO: 8 <u>Protein sequence</u>
Protein Accession #: NP_002416

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     MHLAFLVLLC LPVCSAYPLS GAAKEEDSNK DLAQQYLEKY YNLEKDVKQF RRKDSNLIVK  60
10   KIQGMQKFLG LEVTGKLDTD TLEVMRKPRC GVPDVGHFSS FPGMPKWRKT HLTYRIVNYT  120
     PDLPRDAVDS AIEKALKVWE EVTPLTPSRL YEGEADIMIS PAVKEHGDPY SPDGPGHSLA  180
     HAYPPGPGLY GDIHPDDDEK WTEDASGTNL FLVAAHELGH SLGLFHSANT EALMYPLYNS  240
15   FTELAQFRLS QDDVNGIQSL YGPPPASTEE PLVPTKSVPS GSEMPAKCDP ALSPDAISTL  300
     RGEYLFFKDR YFWRRSHWNP EPEFHLISAF WPSLPSYLDA AYEVNSRDTV FIFKGNEFWA  360
     IRGNEVQAGY PRGIHTLGFP PTIRKIDAAV SDKEKKKTYF FAADKYWRFD ENSQSMEQGF  420
     PRLIADDFPG VSPKVDAVLQ AFGFFYFFSG SSQFEFDPNA RMVTHILKSN SWLHC
```

Seq ID NO: 9 <u>DNA sequence</u>
Nucleic Acid Accession #: XM_058189.2
Coding sequence: 169..774

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     GAAGACCAGC TCAGCTCTTC AGTTGTTGAT CATTGTCTAT TGTCTCCAA ACAGTAAACC   60
     AGTATTTCAC ACTGAGATTG TCGGCTGCGG GTATATTCCA ATTCCCCGTC TCCTCATGAA  120
     TATGAAGTGA AGGGCTCTGA CCCTGGAAGT GGTTCTAAGC GGAGAAATAG GGGGTCTCGG  180
     AAGTGTGGAG GCTGCCTAAG TTGTTTGCTG ATTCCGCTTG CACTTGGAG TATAATCGTG   240
30   AACATATTAT TGTATTTCCC GAATGGGCAA ACTTCCTATG CATCCAGCAA TAAACTCACC  300
     AACTACGTGT GGTATTTTGA AGGAATCTGT TTCTCAGGCA TCATGATGCT TATAGTAACA  360
     ACAGTTCTTC TGGTACTGGA GAATAATAAC AACTATAAAT GTTGCCAGAG TGAAAACTGC  420
     AGCAAAAAAT ATGTGACACT GCTGTCAATT ATCTTTTCTT CCCTCGGAAT TGCTTTTTCT  480
     GGATACTGCC TGGTCATCTC TGCCTTGGGT CTTGTCCAAG GGCCATATTG CCGCACCCTT  540
35   GATGGCTGGG AGTATGCTTT TGAAGGCACT GCTGGACGTT TCCTTACAGA TTCTAGCATA  600
     TGGATTCAGT GCCTGGAACC TGCACATGTT GTGGAGTGGA ACATCATTTT ATTTTCCATT  660
     CTCATAACCC TCAGTGGGCT TCAAGTGATC ATCTGCCTCA TCAGAGTAGT CATGCAACTA  720
     TCCAAGATAC TGTGTGGAAG CTATTCAGTG ATCTTCCAGC CTGGAATCGT TTGAATAAGG  780
     ACAAAATGTT TTCCATTATC AAGACATGGC CATCTATCTA AATATTATAT CAACTGTGTA  840
40   GACTTGAGGG CAATATTGAA ATGATGGTGC TTTCTGCATT TGGTGTTTAT TTGTAAAAAA  900
     TTTGCAGTCC TCACTGCACA TGCAAGTATA CCACCCTTCC ATTTAGTATG TTTTTTAAGT  960
     AATATGCATC AGAAACTTCA GAAATACTTC TGCCCTTTGA TCAAACAAAT CCATTTCCAA  1020
     GAATCTGTAC TAGGGAAGTA AATAAGAATA TGAGAGAAAC CTTTATGCAA ATATGTATAT  1080
     TGCAACATTA TTTAATATTC TGGAAAATTG GAAACACCCC AAAATTCTAA ACTCAGAGGA  1140
45   AGGATTAAGT AAAGAGTGGT ACATACTGTA AATGTTTTCT GATATTAAAA AAAAAATTAA  1200
     ATAAAAAATA AAGAGTACTA CATGGTTGTA AAA
```

Seq ID NO: 10 <u>Protein sequence</u>
Protein Accession #: XP_058189.1

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     MGSRKCGGCL SCLLIPLALW SIIVNILLYF PNGQTSYASS NKLTNYVWYF EGICFSGIMM  60
     LIVTTVLLVL ENNNNYKCCQ SENCSKKYVT LLSIIFSSLG IAPSGYCLVI SALGLVQGPY  120
     CRTLDGNEYA FEGTAGRFLT DSSIWIQCLE PAHVVEWNII LFSILITLSG LQVIICLIRV  180
     VMQLSKILCG SYSVIFQPGI I
```

Seq ID NO: 11 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_002421.2
Coding sequence: 1..1409

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     ATGCACAGCT TTCCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTC ACACAGCTTC   60
     CCAGCGACTC TAGAAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAAA  120
     TACTACAACC TGAAGAATGA TGGGGAGGCAA GTTGAAAAGC GGAGAAATAG TGGCCCAGTG  180
     GTTGAAAAAT TGAAGCAAAT GCAGGAATTC TTTGGGCTGA AAGTGACTGG GAAACCAGAT  240
     GCTGAAACCC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCCTGATGT GGCTCAGTTT  300
70   GTCCTGACTG AGGGGAACCC TCGCTGGGAG CAAACACATC TTGACCTACAG GATTGAAAAT  360
     TACACGCCAG ATTTGCCAAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTTCCAACTC  420
     TGGAGTAATG TCACACCTCT GACATTCACC AAGGTCTCTG AGGGTCAAGC AGACATCATG  480
     ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCCTT TTGATGGACC TGGAGGAAAT  540
     CTTGCTCATG CTTTTCAACC AGGCCCAGGT ATTGGAGGGG ATGCTCATTT TGATGAAGAT  600
75   GAAAGGTGGA CCAACAATTT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGAACTC  660
     GGCCATTCTC TTGGACTCTC CCATTCTACT GATATCGGGG CTTTGATGTA CCCTAGCTAC  720
     ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATTG ATGGCATCCA AGCCATATAT  780
     GGACGTTCCC AAAATCCTGT CCAGCCCATC GGCCCACAAA CCCCAAAAGC ATGTGACAGT  840
     AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA  900
80   TTCTACATGC GCACAAATCC CTTCTACCCG GAAGTTGAGC TCAATTTCAT TTCTGTTTTC  960
     TGGCCACAAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC  1020
     CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC  1080
     CCCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAACTG TGAAGCATAT CGATGCTGCT  1140
     CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTTG CTAACAAATA CTGGAGGTAT  1200
     GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTCCT  1260
```

```
GGAATTGGCC ACAAAGTTGA TGCAGTTTTC ATGAAAGATG GATTTTTCTA TTTCTTTCAT   1320
GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT   1380
AATAGCTGGT TCAACTGCAG GAAAAATTAG
```

Seq ID NO: 12 Protein sequence
Protein Accession #:   NP_002412.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MHSPPPLLLL LFWGVVSHSF PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV   60
VEKLKQMQEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRWE QTHLTYRIEN   120
YTPDLPRADV DHAIEKAPQL WSNVTPLTFT KVSEGQADIM ISFVRGDHRD NSPFDGPGGN   180
LAHAFQPGPG IGGDAHFDED ERWTNNFREY NLHRVAAHEL GHSLGLSHST DIGALMYPSY   240
TFSGDVQLAQ DDIDGIQAIY GRSQNPVQPI GPQTPKACDS KLTFDAITTI RGEVMFFKDR   300
FYMRTNPFYP EVELNFISVF WPQLPNGLEA AYEFADRDEV RPFKGNKYWA VQGQNVLHGY   360
PKDIYSSFGF PRTVKHIDAA LSEENTGKTY FFVANKYWRY DEYKRSMDPG YPKMIAHDFP   420
GIGHKVDAVF MKDGFFYFFH GTRQYKFDPK TKRILTLQKA NSWFNCRKN
```

Seq ID NO: 13 DNA sequence
Nucleic Acid Accession #: NM_002421.2
Coding sequence: 1..1409

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCACAGCT TTCCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTC ACACAGCTTC   60
CCAGCGACTC TAGAAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAAA   120
TACTACAACC TGAAGAATGA TGGGAGGCAA GTTGAAAAGC GGAGAAATAG TGGCCCAGTG   180
GTTGAAAAAT TGAAGCAAAT GCAGGAATTC TTTGGGCTGA AAGTGACTGG GAAACCAGAT   240
GCTGAAACCC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCCTGATGT GGCTCAGTTT   300
GTCCTCACTG AGGGGAACCC TCGCTGGGAG CAAACACATC TGACCTACAG GATTGAAAAT   360
TACACGCCAG ATTTGCCAAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTTCCAACTG   420
TGGAGTAATG TCACACCTCT GACATTCACC AAGGTCTCTG AGGGTCAAGC AGACATCATG   480
ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCCTT TTGATGGACC TGGAGGAAAT   540
CTTGCTCATG CTTTTCAACC AGGCCCAGGT ATTGGAGGGG ATGCTCATTT TGATGAAGAT   600
GAAAGGTGGA CCAACAATTT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGCCCTC   660
GGCCATTCTC TTGGACTCTC CCATTCTACT GATATCGGGG CTTTGATGTA CCCTAGCTAC   720
ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATTG ATGGCATCCA AGCCATATAT   780
GGACGTTCCC AAAATCCTGT CCAGCCCATC GGCCCAACAA CCCCAAAAGC ATGTGACAGT   840
AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA   900
TTCTACATGC GCACAAATCC CTTCTACCCG GAAGTTGAGC TCAATTTCAT TTCTGTTTTC   960
TGGCCACAAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC   1020
CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC   1080
CCCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAACTG TGAAGCATAT CGATGCTGCT   1140
CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTTG CTAACAAATA CTGGAGGTAT   1200
GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTCCT   1260
GGAATTGGCC ACAAAGTTGA TGCAGTTTTC ATGAAAGATG GATTTTTCTA TTTCTTTCAT   1320
GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT   1380
AATAGCTGGT TCAACTGCAG GAAAAATTAG
```

Seq ID NO: 14 Protein sequence
Protein Accession #:   NP_002412.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MHSPPPLLLL LFWGVVSHSF PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV   60
VEKLKQMQEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRWE QTHLTYRIEN   120
YTPDLPRADV DHAIEKAPQL WSNVTPLTFT KVSEGQADIM ISFVRGDHRD NSPFDGPGGN   180
LAHAFQPGPG IGGDAHFDED ERWTNNFREY NLHRVAAHAL GHSLGLSHST DIGALMYPSY   240
TFSGDVQLAQ DDIDGIQAIY GRSQNPVQPI GPQTPKACDS KLTFDAITTI RGEVMFFKDR   300
FYMRTNPFYP EVELNFISVF WPQLPNGLEA AYEFADRDEV RPFKGNKYWA VQGQNVLHGY   360
PKDIYSSFGF PRTVKHIDAA LSEENTGKTY FFVANKYWRY DEYKRSMDPG YPKMIAHDFP   420
GIGHKVDAVF MKDGFFYFFH GTRQYKFDPK TKRILTLQKA NSWFNCRKN
```

Seq ID NO: 15 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 141..1580

```
1          11         21         31         41         51
|          |          |          |          |          |
TCTGCGTGTG CCGGGGCTAG GGGCTGGAAG TCCTGGCTCT AGTTGCACCT CGGAAGGAAA   60
AGGCAAACAG AGGAGGGGAG GCGTCTTAGG ACTGCCTGGA TCCAGAGCAC TTTCCTCGGC   120
CTCTACAGGC CTGTGTCGCT ATGGGTTCCC CCGCCGCCCC GGAGGGAGCG CTGGGCTACG   180
TCCGCGAGTT CACTCGCCAC TCCTCCGACG TGCTGGGCAA CCTCAACGAG CTGCGCCTGC   240
GCGGGATCCT CACTGACGTC ACGCTGCTGG TTGGCGGGCA ACCCTCAGA GCACACAGG   300
CAGTTCTCAT CGCCTGCAGT GGCTTCTTCT ATTCAATTTT CCGGGGCCGT GCGGGAGTCG   360
GGGTGGACGT GCTCTCTCTG CCCGGGGGTC CCGAAGCGAG AGGCTTCGCC CCTCTATTGG   420
ACTTCATGTA CACTTCGCGC CTGCGCCTCT CTCCAGCCAC TGCACCAGCA GTCCTAGCGG   480
CCGCCACCTA TTTGCAGATG GAGCACGTGG TCCAGGCATG CCACCGCTTC ATCCAGGCCA   540
GCTATGAACC TCTGGGCATC TCCCTGCGCC CCCTGGAAGC AGAACCCCCA ACACCCCCAA   600
CGGCCCCTCC ACCAGGTAGT CCCAGGCGCT CGGAAGGACA CCCAGACCCA CCTACTGAAT   660
CTCGAAGCTG CAGTCAAGGC CCCCCCAGTC CAGCCAGCCC TGACCCCAAG GCCTGCAACT   720
GGAAAAAGTA CAAGTACATC GTGCTAAACT CTCAGGCCTC CCAAGCAGGG AGCCTGGTCG   780
```

```
GGGAGAGAAG TTCTGGTCAA CCTTGCCCCC AAGCCAGGCT CCCCAGTGGA GACGAGGCCT    840
CCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GTGAAGAAGG ACCCATTCCT GGTCCCCAGA    900
GCAGGCTCTC TCCAACTGCT GCCACTGTGC AGTTCAAATG TGGGGCTCCA GCCAGTACCC    960
CCTACCTCCT CACATCCCAG GCTCAAGACA CCTCTGGATC ACCCTCTGAA CGGGCTCGTC   1020
CACTACCGGG AAGTGAATTT TTCAGCTGCC AGAACTGTGA GGCTGTGGCA GGGTGCTCAT   1080
CGGGGCTGGA CTCCTTGGTT CCTGGGGACG AAGACAAACC CTATAAGTGT CAGCTGTGCC   1140
GGTCTTCGTT CCGCTACAAG GGCAACCTTG CCAGTCATCG TACAGTGCAC ACAGGGGAAA   1200
AGCCTTACCA CTGCTCAATC TGCGGAGCCC GTTTTAACCG GCCAGCAAAC CTGAAAACGC   1260
ACAGCCGCAT CCATTCGGGA GAGAAGCCGT ATAAGTGTGA GACGTGCGGC TCGCGCTTTG   1320
TACAGGTGGC ACATCTGCGG GCGCACGTGC TGATCCACAC CGGGGAGAAG CCCTACCCTT   1380
GCCCTACCTG CGGAACCCGC TTCCGCCACC TGCAGACCCT CAAGAGCCAC GTTCGCATCC   1440
ACACCGGAGA GAAGCCTTAC CACTGCGACC CCTGTGGCCT GCATTCCGG CACAAGAGTC    1500
AACTGCGGCT GCATCTGCGC CAGAAACACG GAGCTGCTAC CAACACCAAA GTGCACTACC   1560
ACATTCTCGG GGGGCCCTAG CTGAGCGCAG GCCCAGGCCC CACTTGCTTC CTGCGGGTGG   1620
GAAAGCTGCA GGCCCAGGCC TTGCTTCCCT ATCAGGCTTG GGCATAGGGG TGTGCCAGGC   1680
CACTTTGGTA TCAGAAATTG CCACCCTCTT AATTTCTCAC TGGGGAGAGC AGGGGTGGCA   1740
GATCCTGGCT AGATCTGCCT CTGTTTTGCT GGTCAAAACC TCTTCCCCAC AAGCCAGATT   1800
GTTTCTGAGG AGAGAGCTAG CTAGGGGCTG GGAAAGGGGA GAGATTGGAG TCCTGGTCTC   1860
CCTAAGGGAA TAGCCCTCCA CCTGTGGCCC CCATTGCATT CAGTTTATCT GTAAATATAA   1920
TTTATTGAGG CCTTTGGGTG GCACCGGGGC CTTCATTCGA TTGCATTTCC CACTCCCCTC   1980
TTCCACAAGT GTGATTAAAA GTGACCAGAA ACACAGAAGG TGAGATCACA GCTCTGCTGG   2040
CAGAGATTAC TAGCCCTTGG CTCTCTCGTT TGGCTTGGGT ATTTTATATT ATTTCTGTCA   2100
TAACTTTTAT CTTTAGAATT GTTCTTTCTC CTGTTTGTTT GCTTGTTAGT TTGTTTAAAA   2160
TGGAAAAAGG GGTTCTCTGT GTTCTGCCCC TGTAATTCTA GGTCTGGAAC CTTTATTTGT   2220
TCTAGGGCAG CTCTGGGAAC ATGCGGGATT GTGGAATTGG GTCAGGAACC CTCTCTGGTA   2280
TTCTGGATGT TGTAGGTTCT CTAGCAGTCT AGAAATGGAT ACAGACATTT CTCTGTTCTT   2340
CAAGGGTGAT AGGAACCATT ATGTTGAGCC CAAAATGGAA GTAATAATAA ATGCCTCCTG   2400
GAGGCTGTGG GTGTGGGGGA TTCTGTATCT GGATTCCGTA TCACTCCAAC TGGAGGGCTGT   2460
GGGTGTGGGG GATTCTGTAT CTGGATTCCG TATCACTCCA AGTGGAGGCT GGCAGGTTTT   2520
TCTGCAAGAT GGTCCAGAAT CTAAAATGTC CCATTAATCT GGTCACTTGG GTTTGGCTCT   2580
GCTGTATCCA TCTATAGTGG TAGAGACCCA CCAGGGCTCA AGTGGAGTCC ATCATCCTCC   2640
CACGGGGGCC TGTTCTTAGC ACTGAGTTGA TCGCTCCATG GGGGAGAGAT CAGACATTCC   2700
TTATCAGAGA TGATGTGACC TTTTCTGACT CTGCCCAGTC TCTATGAATG TTATGGCCTA   2760
GGGAAGAAATC ATGAAACTCT TTAGCTTGAT TAGATGGTAA ACAGTGTTAA CCCATCCTTT   2820
ACTACAGAGG CATATGGGTT TGAATGTTAC CTGGGGTTCT CTCTATTGAG TTGACCCCCT   2880
TCTTCCTTTA GTGGGTTTTG GACATCTTCT GGCAAGTGTC CAGATGCCAG AACCTTCTTT   2940
TCCTCTAGAA GGGATGGTGC TTGGTAACCT TACCTTTTAA AAGCTGGGTC TGTGACCTGG   3000
TCTTCCCATC CCTGCATTCC TGTCTGGAAC CAGTGAATGC ATTAGAACCT TCCATAGGAA   3060
AAGAAAAGGG GCTGAGTTCC ATTCTGGGTT TGCTGTAGTT TGGTTCGGAT TATTGTTGGC   3120
ATTACAGATG TAAAAGATTG ACTAGCCCAT AGGCCAAAGG CCTGTTCTAG TTGACCAAGT   3180
TTCAAGTAGG ATTAAGAGGT TGGTTGAGGG GTGCAGTTTC TGGTGTAGGC CAGGTAGGTA   3240
GAAAGTGAGG AACAGGGTTG CCTCTTGGCT GGGTGGAGTC TCTGAAATGT TAGAAGAAGC   3300
GCTGAAGCCT TGATTGATAG TTCTGCCCCT TGTTGCCCTG GGGCTTATCT GATTATGGGA   3360
CGAGGGTAGA AAGTAAGAAG CACTTTTGAA TTTGTGGGGT AGAACTTCAA CAATAAGTCA   3420
GTTCTAGTGG CTGTCGCCTG GGGACTAGTG AGAAAGCTAC TCTTCTCCCT CTTCCCTCTT   3480
TCTCCCCATG GCCCCACTGC AGAATTAAAG AAGGAAGAAG GGAAGGCGGA GGAGTCTATA   3540
AGAAGGAATC ATGATTTCTA TTTAGCAGAT TGGATGGGCA GGTGGAGAAT GCCTGGGGGT   3600
AGAAATGTTA GATCTTGCAA CATCAGATCC TTGGAATAAA GAAGCCTCTC TGYGCWRAAA   3660
AAAAAAAAAA AAAAAA
```

Seq ID NO: 16 Protein sequence
Protein Accession #: FGENESH predicted

```
1          11         21         31         41         51
|          |          |          |          |          |
MGSPAAPEGA LGYVREFTRH SSDVLGNLNB LRLRGILTDV TLLVGGQPLR AHKAVLIACS     60
GPFYSIPRGR AGVGVDVLSL PGGPEARGFA PLLDFMYTSR LRLSPATAPA VLAAATYLQM    120
EHVVQACHRF IQASYEPLGI SLRPLEAEPP TPPTAPPPGS PRRSEGHPDP PTESRSCSQG    180
PPSPASFDPK ACNWKKYKYI VLNSQASQAG SLVGERSSGQ PCFQARLPSG DEASSSSSSS    240
SSSSEEGPIP GPQSRLSPTA ATVQFKCGAP ASTPYLLTSQ ADTSGSPSE RARPLPGSEF    300
FSCQNCEAVA GCSSGLDSLV PGDEDKPYKC QLCRSSFRYK GNLASHRTVH TGEKPYHCSI    360
CGARFNRPAN LKTHSRIHSG EKPYKCETCG SRFVQVAHLR AEVLIHTGEK PYPCPTCGTR    420
FRHLQTLKSH VRIHTGEKPY HCDPCGLHFR HKSQLRLHLR QKHGAATNTK VHYHILGGP
```

Seq ID NO: 17 DNA sequence
Nucleic Acid Accession #: XM_039209
Coding sequence: 1..2049

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCTGAAGA TGCTCTCCTT TAAGCTGCTG CTGCTGGCCG TGGCTCTGGG CTTCTTTGAA     60
GGAGATGCTA AGTTTGGGGA AAGAAACGAA GGGAGCGGAG CAAGGAGGAG AAGGTGCCTG    120
AATGGGAACC CCCGAAGCG CCTGAAAAGG AGAGACAGGA GGATGATGTC CCAGCTGGAG     180
CTGCTGAGTG GGGGAGAGAT GCTGTGCGGT GGCTTCTACC CTCGGCTGTC CTGCTGCCTG    240
CGGAGTGACA GCCCGGGGCT AGGGCGCCTG GAGAATAAGA TATTTCTGT TACCAACAAC     300
ACAGAATGTG GGAAGTTACT GGAGGAAATC AAATGTGCAC TTTGCTCTCC ACATTCTCAA    360
AGCCTGTTCC ACTCACCTGA GAGAGAAGTC TTGGAAAGAG ACCTAGTACT TCCTCTGCTC    420
TGCAAAGACT ATTGCAAAGA ATTCTTTTAC ACTTGCCGAG GCCATATTCC AGGTTTCCTT    480
CAAACAACTG CGGATGAGTT TTGCTTTTAC TATGCAAGAA AAGATGGTGG GTTGTGCTTT    540
CCAGATTTTC CAAGAAAACA AGTCAGAGGA CCAGCATCTA ACTACTTGGA CCAGATGGAA    600
GAATATGACA AAGTGGAAGA GATCAGCAGA AAGCACAAAC ACAACTGCTT CTGTATTCAG    660
GAGGTTGTGA GTGGGCTGCG GCAGCCCGTT GGTGCCCTGC ATAGTGGGGA TGGCTCGCAA    720
CGTCTCTTCA TTCTGGAAAA AGAAGGTTAT GTGAAGATAC TTACCCCTGA AGGAGAAATT    780
TTCAAGGAGC CTTATTTGGA CATTCACAAA CTTGTTCAAA GTGGAATAAA GGGAGGAGAT    840
```

```
     GAAAGAGGAC TGCTAAGCCT CGCATTCCAT CCCAATTACA AGAAAAATGG AAAGTTGTAT    900
     GTGTCCTATA CCACCAACCA AGAACGGTGG GCTATCGGGC CTCATGACCA CATTCTTAGG    960
     GTTGTGGAAT ACACAGTATC CAGAAAAAAT CCACACCAAG TTGATTTGAG AACAGCCAGA   1020
     GTCTTTCTTG AAGTTGCAGA ACTCCACAGA AAGCATCTGG GAGGACAACT GCTCTTTGGC   1080
     CCTGACGGCT TTTTGTACAT CATTCTTGGT GATGGGATGA TTACACTGGA TGATATGGAA   1140
     GAAATGGATG GGTTAAGTGA TTTCACAGGC TCAGTGCTAC GGCTGGATGT GGACACAGAC   1200
     ATGTGTAACG TGCCTTATTC CATACCAGG AGCAACCCAC ACTTCAACAG CACCAACCAG    1260
     CCCCCCGAAG TGTTTGCTCA TGGGCTCCAC GATCCAGGCA GATGTGCTGT GGATAGACAT   1320
     CCCACTGATA TAAACATCAA TTTAACGATA CTGTGTTCAG ACTCCAATGG AAAAAACAGA   1380
     TCATCAGCCA GAATTCTACA GATAATAAAG GGGAAAGATT ATGAAAGTGA GCCATCACTT   1440
     TTAGAATTCA AGCCATTCAG TAATGGTCCT TTGGTTGGTG GATTTGTATA CCGGGGCTGC   1500
     CAGTCAGAAA GATTGTATGG AAGCTACGTG TTTGGAGATC GTAATGGGAA TTTCCTAACT   1560
     CTCCAGCAAA GTCCTGTGAC AAAGCAGTGG CAAGAAAAAC CACTCTGTCT CGGCACTAGT   1620
     GGGTCCTGTA GAGGCTACTT TTCCGGTCAC ATCTTGGGAT TTGGAGAAGA TGAACTAGGT   1680
     GAAGTTTACA TTTTATCAAG CAGTAAAAGT ATGACCCAGA CTCACAATGG AAAACTCTAC   1740
     AAAATTGTAG ATCCCAAAAG ACCTTTAATG CCTGAGGAAT GCAGAGCCAC GGTACAACCT   1800
     GCACAGACAC TGACTTCAGA GTGCTCCAGG CTCTGTCGAA ACGGCTACTG CACCCCCACG   1860
     GGAAAGTGCT GCTGCAGTCC AGGCTGGGAG GGGGACTTCT GCAGAACTGC AAAATGTGAG   1920
     CCAGCATGTC GTCATGGAGG TGTCTGTGTT AGACCGAACA AGTGCCTCTG TAAAAAAGGA   1980
     TATCTTGGTC CTCAATGTGA ACAAGTGGAC AGAAACATCC GCAGAGTGAC CAGGGCAGAC   2040
     ATCACCTAG
```

Seq ID NO: 18 Protein sequence
Protein Accession #: XP_039209

```
      1          11         21         31         41         51
      |          |          |          |          |          |
     MLKMLSPKLL LLAVALGFFE GDAKFGERNE GSGARRRRCL NGNPPKRLKR RDRRMMSQLE    60
     LLSGGEMLCG GFYPRLSCCL RSDSPGLGRL ENKIFSVTNN TECGKLLEEI KCALCSPHSQ   120
     SLFHSPEREV LERDLVLPLL CKDYCKEFFY TCRGHIPGFL QTTADSFCFY YARKDGGLCF   180
     PDFPRKQVRG PASNYLDQME EYDKVEEISR KHKHNCFCIQ EVVSGLRQPV GALHSGDGSQ   240
     RLFILEKEGY VKILTPEGEI FKEPYLDIHK LVQSGIKGGD ERGLLSLAFH PNYKKNGKLY   300
     VSYTTNQERW AIGPHDHILR VVEYTVSRKN PHQVDLRTAR VFLEVAELHR KHLGGQLLPG   360
     PDGFLYIILG DGMITLDDME EMDGLSDFTG SVLRLDVDTD MCNVPYSIPR SNPHFNSTNQ   420
     PPPVFAHGLH DPGRCAVDRH PTDININLTI LCSDSNGKNR SSARIIQIIK GKDYESEPSL   480
     LEFKPFSNGP LVGGFVYRGC QSERLYGSYV FGDRNGNFLT LQQSPVTKQW QEKPLCLGTS   540
     GSCRGYFSGH ILGFGEDELG EVYILSSSKS MTQTHNGKLY KIVDPKRPLM PEECRATVQP   600
     AQTLTSECSR LCRNGYCTPT GKCCCSPGWE GDFCRTAKCE PACRHGGVCV RPNKCLCKKG   660
     YLGPQCEQVD RNIRRVTRAD IT
```

Seq ID NO: 19 DNA sequence
Nucleic Acid Accession #: NM_014331.2
Coding sequence: 1..1506

```
      1          11         21         31         41         51
      |          |          |          |          |          |
     ATGGTCAGAA AGCCTGTTGT GTCCACCATC TCCAAAGGAG GTTACCTGCA GGGAAATGTT     60
     AACGGGAGGC TGCCTTCCCT GGGCAACAAG GAGCCACCTG GGCAGGAGAA AGTGCAGCTG    120
     AAGAGGAAAG TCACTTTACT GAGGGGAGTC TCCATTATCA TTGGCACCAT CATTGGAGCA    180
     GGAATCTTCA TCTCTCCTAA GGGCGTGCTC CAGAACACGG GCAGCGTGGG CATGTCTCTG    240
     ACCATCTGGA CGGTGTGTGG GGTCCTGTCA CTATTTGGAG CTTTGTCTTA TGCTGAATTG    300
     GGAACAACTA TAAAGAAATC TGGAGGTCAT TACACATATA TTTTGGAAGT CTTTGGTCCA    360
     TTACCAGCTT TTGTACGAGT CTGGGTGGAA CTCCTCATAA TACGCCCTGC AGCTACTGCT    420
     GTGATATCCC TGGCATTTGG ACGCTACATT CTGGAACCAT TTTTATTCA ATGTGAAATC     480
     CCTGAACTTG CGATCAAGCT CATTACAGCT GTGGGCATAA CTGTAGTGAT GGTCCTAAAT    540
     AGCATGAGTG TCAGCTGGAG CGCCCGGATC CAGATTTTCT TAACCTTTTG CAAGCTCACA    600
     GCAATTCTGA TAATTATAGT CCCTGGAGTT ATGCAGCTAA TTAAAGGTCA ACGCAGAAC     660
     TTTAAAGACG CGTTTTCAGG AAGAGATTCA AGTATTACGC GGTTGCCACT GGCTTTTTAT    720
     TATGAAATGT ATGCATATGC TGGCTGGTTT TACCTCAACT TTGTTACTGA AGAAGTAGAA    780
     AACCCTGAAA AAACCATTCC CCTTGCAATA TGTATATCCA TGGCCATTGT CACCCATTGGC   840
     TATGTGCTGA CAAATGTGGC CTACTTTACG ACCATTAATG CTGAGGAGCT GCTGCTTTCA    900
     AATGCAGTGG CAGTGACCTT TTCTGAGCGG CTACTGGGAA ATTTCTCATT AGCAGTTCCG    960
     ATCTTTGTTG CCCTCTCCTG CTTTGGCTCC ATGAACGGTG GTGTGTTTGC TGTCTCCAGG   1020
     TTATTCTATG TTGCGTCTCG AGAGGGTCAC CTTCCAGAAA TCCTCTCCAT GATTCATGTC   1080
     CGCAAGCACA CTCCTCTACC AGCTGTTATT GTTTTGCACC CTTTGACAAT GATAATGCTC   1140
     TTCTCTGGAG ACCTCGACAG TCTTTTGAAT TTCCTCAGTT TTGCCAGGTG GCTTTTTATT   1200
     GGGCTGGCAG TTGCTGGGCT GATTTATCTT CGATCAAAT GCCCAGATAT GCATCGTCCT    1260
     TTCAAGGTGC CACTGTTCAT CCCAGCTTTG TTTTCCTTCA CATGCCTCTT CATGGTTCCC   1320
     CTTTCCCTCT ATTCGGACCC ATTTAGTACA GGGATTGGCT TCGTCATCAC TCTGACTGGA   1380
     GTCCCTGCGT ATTATCTCTT TATTATATGG GACAAGAAAC CCAGGTGGTT TAGAATAATG   1440
     TCAGAGAAAA TAACCAGAAC ATTACAAATA ATACTGGAAG TTGTACCAGA AGAAGATAAG   1500
     TTATGAACTA ATGGACTTGA GATCTTGGCA ATCTGCCCAA GGGGAGACAC AAAATAGGGA   1560
     TTTTTACTTC ATTTTCTGAA AGTCTAGAGA ATTACAACTT TGGTGATAAA CAAAGGGAGT   1620
     CAGTTATTTT TATTCATATA TTTTAGCATA TTCGAACTAA TTTCTAAGAA ATTTAGTTAT   1680
     AACTCTTATGT AGTTATAGAA AGTGAAATATG CAGTTATTCT ATGAGTCGCA CAATTCTTGA   1740
     GTCTCTGATA CCTACCTATT GGGGTTAGGA GAAAAGACTA GACAATTACT ATGTGGTCAT   1800
     TCTCTACAAC ATATGTTAGC ACGGCAAAGA ACCTTCAAAT TGAAGACTGA GATTTTTCTG   1860
     TATATATGGG TTTTGTAAAG ATGGTTTTAC ACACTACAGA TGTCTATACT GTGAAAAGTG   1920
     TTTTCAATTC TGAAAAAAAG CATACATCAT GATTATGGCA AAGAGGAGAG AAAGAAATTT   1980
     ATTTTACATT GACATTGCAT TGCTTCCCCT TAGATACCAA TTTAGATAAC AAACACTCAT   2040
     GCTTTAATGG ATTATACCCA GAGCCACTTTG AACAAAGGTC AGTGGGGATT GTTGAAATACA  2100
     TTAAAGAAGA GTTTCTAGGG GCTACTGTTT ATGAGACACA TCCAGGAGTT ATGTTTAAGT   2160
     AAAAATCCTT GAGAATTTAT TATGTCAGAT GTTTTTTCAT TCATTATCAG GAAGTTTTAG   2220
```

```
TTATCTGTCA TTTTTTTTTT TCACATCAGT TTGATCAGGA AAGTGTATAA CACATCTTAG   2280
AGCAAGAGTT AGTTTGGTAT TAAATCCTCA TTAGAACAAC CACCTGTTTC ACTAATAACT   2340
TACCCCTGAT GAGTCTATCT AAACATATGC ATTTTAAGCC TTCAAATTAC ATTATCAACA   2400
TGAGAGAAAT AACCAACAAA GAAGATGTTC AAAATAATAG TCCCATATCT GTAATCATAT   2460
CTACATGCAA TGTTAGTAAT TCTGAAGTTT TTTAAATTTA TGGCTATTTT TACACGATGA   2520
TGAATTTTGA CAGTTTGTGC ATTTTCTTTA TACATTTTAT ATTCTTCTGT TAAAATATCT   2580
CTTCAGATGA AACTGTCCAG ATTAATTAGG AAAAGGCATA TATTAACATA AAAATTGCAA   2640
AAGAAATGTC GCTGTAAATA AGATTTACAA CTGATGTTTC TAGAAAATTT CCACTTCTAT   2700
ATCTAGGCTT TGTCAGTAAT TTCCACACCT TAATTATCAT TCAACTTGCA AAAGAGACAA   2760
CTGATAAGAA GAAAATTGAA ATGAGAATCT GTGGATAAGT GTTTGTGTTC AGAAGATGTT   2820
GTTTTGCCAG TATTAGAAAA TACTGTGAGC CGGGCATGGT GGCTTACACTC TGTAATCCCA   2880
GCACTTTGGG AGGCTGAGGG GGTGGATCAC CTGAGGTCGG GAGTTCTAGA CCAGCCTGAC   2940
CAACATGGAG AAACCCCATC TCTACTAAAA ATACAAAATT AGCTGGGCAT GGTGGCACAT   3000
GCTGGTAATC TCAGCTATTG AGGAGGCTGA GGCAGGAGAA TTGCTTGAAC CCGGGAGGCG   3060
GAGGTTGCAG TGAGCCAAGA TTGCACCACT GTACTCCAGC CTGGGTGACA AAGTCAGACT   3120
CCATCTCCAA AAAAAAAAAA AAAA
```

Seq ID NO: 20 <u>Protein sequence</u>
Protein Accession #: NP_055146.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MVRKPVVSTI SKGGYLQGNV NGRLPSLGNK EPPGQEKVQL KRKVTLLRGV SIIIGTIIGA    60
GIFISPKGVL QNTGSVGMSL TIWTVCGVLS LFGALSYAEL GTTIKKSGGH YTYILEVFGP   120
LPAFVRVWVE ILIIRPAATA VISLAFGRYI LEPFFIQCEI PELAIKLITA VGITVVMVLN   180
SMSVSWSARI QIFLTPCKLT AILIIIVPGV MQLIKGGTQN FKDAPSGRDS SITRLPLAFY   240
YGMYAYAGWF YLNFVTEEVE NPEKTIPLAI CISMAITIGV YVLTNVAYFT TINAEELLLS   300
NAVAVTFSER LLGNFSLAVP IPVALSCFGS MNGGVPAVSR LFYVASREGH LPEILSMIHV   360
RKHTPLPAVI VLRPLTMIML FSGDLDSLLN FLSFARWLFI GLAVAGLIYL RYKCPDMHRP   420
PKVPLFIPAL FSPTCLFMVA LSLYSDPFST GIGFVITLTG VPAYYLFIIW DKKPRWFRIM   480
SEKITRTLQI ILEVVPEEDK L
```

Seq ID NO: 21 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_002422.2
Coding sequence: 64..1497

```
1          11         21         31         41         51
|          |          |          |          |          |
ACAAGGAGGC AGGCAAGACA GCAAGGCATA GAGACAACAT AGAGCTAAGT AAAGCCAGTG    60
GAAATGAAGA GTCTTCCAAT CCTACTGTTG CTGTGCGTGG CAGTTTGCTC AGCCTATCCA   120
TTGGATGGAG CTGCAAGGGG TGAGGACACC AGCATGAACC TTGTTCAGAA ATATCTAGAA   180
AACTACTACG ACCTCAAAAA AGATGTGAAA CAGTTTGTTA GGAGAAAGGA CAGTGGTCCT   240
GTTGTTAAAA AAATCCGAGA AATGCAGAAG TTCCTTGGAT TGGAGGTGAC GGGGAAGCTG   300
GACTCCGACA CTCTGGAGGT GATGCGCAAG CCCAGGTGTG GAGTTCCTGA TGTTGGTCAC   360
TTCAGACCT TTCCTGGCAT CCCGAAGTGG AGGAAAACCC ACCTTACATA CAGGATTGTG   420
AATTATACAC CAGATTTGCC AAAAGATGCT GTTGATTCTG CTGTTGAGAA AGCTCTGAAA   480
GTCTGGGAAG AGGTGACTCC ACTCACATTC TCCAGGCTGT ATGAAGGAGA GGCTGATATA   540
ATGATCTCTT TTGCAGTTAG AGAACATGGA GACTTTTACC CTTTTGATGG ACCTGGGAAT   600
GTTTTGGCCC ATGCCTATGC CCCTGGGCCA GGGATTAATG GAGATGCCCA CTTTGATGAT   660
GATGAACAAT GGACAAAGGA TACAACAGGG ACCAATTTAT TTCTCGTTGC TGCTCATGAA   720
ATTGGCCACT CCCTGGGTCT CTTTCACTCA GCCAACACTG AAGCTTTGAT GTACCCACTC   780
TATCACTCAC TCACAGACCT GACTCGGTTC CGCCTGTCTC AAGATGATAT AAATGGCATT   840
CAGTCCCTCT ATGGACCTCC CCCTGACTCC CCTGAGACCC CCCTGGTACC CACGGAACCT   900
GTCCCTCCAG AACCTGGGAC GCCAGCCAAC TGTGATCCTG CTTTGTCCTT TGATGCTGTC   960
AGCACTCTGA GGGGAGAAAT CCTGATCTTT AAAGACAGGC ACTTTTGGCG CAAATCCCTC  1020
AGGAAGCTTG AACCTGAATT GCATTTGATC TCTTCATTTT GGCCATCTCT TCCTTCAGGC  1080
GTGGATGCCG CATATGAAGT TACTAGCAAG GACCTCGTTT TCATTTTTAA AGGAAATCAA  1140
TTCTGGGCCA TCAGAGGAAA TGAGGTACGA GCTGGATACC CAAGAGGCAT CCACACCCTA  1200
GGTTTCCCTC CAACCGTGAG GAAAATCGAT GCAGCCATTT CTGATAAGGA AAAGAACAAA  1260
ACATATTTCT TTGTAGAGGA CAAATACTGG AGATTTGATG AGAAGAGAAA TTCCATGGAG  1320
CCAGGCTTTC CCAAGCAAAT AGCTGAAGAC TTTCCAGGGA TTGACTCAAA GATTGATGCT  1380
GTTTTTGAAG AATTTGGGTT CTTTTATTTC TTTACTGGAT CTTCACAGTT GGAGTTTGAC  1440
CCAAATGCAA AGAAAGTGAC ACACACTTTG AAGAGTAACA GCTGGCTTAA TTGTTGAAAG  1500
AGATATGTAG AAGGCACAAT ATGGGCACTT TAAATGAAGC TAATAATTCT TCACCTAAGT  1560
CTCTGTGAAT TGAAATGTTC GTTTTCTCCT GCCTGTGCTG TGACTCGAGT CACACTCAAG  1620
GGAACTTGAG CGTGAATCTG TATCTTGCCG GTCATTTTTA TGTTATTACA GGGCATTCAA  1680
ATGGGCTGCT GCTTAGCTTG CACCTTGTCA CATAGAGTGA TCTTTCCCAA GAGAAGGGGA  1740
AGCACTCGTG TGCAACAGAC AAGTGACTGT ATCTGTGTAG ACTATTTGCT TATTAATAA   1800
AGACGATTTG TCAGTTGTTT T
```

Seq ID NO: 22 <u>Protein sequence</u>
Protein Accession #: NP_002413

```
1          11         21         31         41         51
|          |          |          |          |          |
MKSLPILLLL CVAVCSAYPL DQAARGEDTS MNLVQKYLEN YYDLEKDVKQ FVRRKDSGPV    60
VKKIREMQKF LGLEVTGKLD SDTLEVMRKP RCGVPDVGHF RTFPGIPKWR KTHLTYRIVN   120
YTPDLPKDAV DSAVEKALKV WEEVTPLTFS RLYEGEADIM ISFAVREHGD FYPFDGPGNV   180
LAHAYAPGPG INGDAHFDDD EQWTKDTTGT NLFLVAAHEI GHSLGLFHSA NTEALMYPLY   240
HSLTDLTRFR LSQDDINGIQ SLYGPPPDSP ETPLVFTEPV PPEPGTPANC DPALSFDAVS   300
TLRGEILIFK DRHFWRKSLR KLEPELHLIS SFWPSLPSGV DAAYEVTSKD LVFIFKGNQF   360
WAIRGNEVRA GYPRGIHTLG FPPTVRKIDA AISDKEKNKT YFFVEDKYWR FDEKRNSMEP   420
GFPKQIAEDF PGIDSKIDAV PEEFGFFYFP TGSSQLEFDP NAKKVTHTLK SNSWLNC
```

Seq ID NO: 23 DNA sequence
Nucleic Acid Accession #: NM_006528
Coding sequence: 57..764

```
1          11         21         31         41         51
|          |          |          |          |          |
GCCGCCAGCG GCTTTCTCGG ACGCCTTGCC CAGCGGGCCG CCCGACCCCC TGCACCATGG    60
ACCCCGCTCG CCCCCTGGGG CTGTCGATTC TGCTGCTTTT CCTGACGGAG GCTGCACTGG   120
GCGATGCTGC TCAGGAGCCA ACAGGAAATA ACGCGGAGAT CTGTCTCCTG CCCCTAGACT   180
ACGGACCCTG CCGGGCCCTA CTTCTCCGTT ACTACTACGA CAGGTACACG CAGAGCTGCC   240
GCCAGTTCCT GTACGGGGGC TGCGAGGGCA ACGCCAACAA TTTCTACACC TGGGAGGCTT   300
GCGACGATGC TTGCTGGAGG ATAGAAAAAG TTCCCAAAGT TTGCCGGCTG CAAGTGAGTG   360
TGGACGACCA GTGTGAGGGG TCCACAGAAA AGTATTTCTT TAATCTAAGT TCCATGACAT   420
GTGAAAAATT CTTTTCCGGT GGGTGTCACC GGAACCGGAT TGAGAACGAG TTTCCAGATG   480
AAGCTACTTG TATGGGCTTC TGCGCACCAA AGAAAATTCC ATCATTTTGC TACAGTCCAA   540
AAGATGAGGG ACTGTGCTCT GCCAATGTGA CTCGCTATTA TTTTAATCCA AGATACAGAA   600
CCTGTGATGC TTTCACCTAT ACTGGCTGTG GAGGGAATGA CAATAACTTT GTTAGCAGGG   660
AGGATTGCAA ACGTGCATGT GCAAAAGCTT TGAAAAAGAA AAAGAAGATG CCAAAGCTTC   720
GCTTTGCCAG TAGAATCCGG AAAATTCGGA AGAAGCAATT TTAAACATTC TTAATATGTC   780
ATCTTGTTTG TCTTTATGGC TTATTTGCCT TTATGGTTGT ATCTGAAGAA TAATATGACA   840
GCATGAGGAA ACAAATCATT GGTGATTTAT TCACCAGTTT TTATTAATAC AAGTCACTTT   900
TTCAAAAATT TGGATTTTTT TATATATAAC TAGCTGCTAT TCAAATGTGA GTCTACCATT   960
TTTAATTTAT GGTTCAACTG TTTGTGAGAC GAATTCTTGC AATGCATAAG ATATAAAAGC  1020
AAATATGACT CACTCATTTC TTGGGGTCGT ATTCCTGATT TCAGAAGAGG ATCATAACTG  1080
AAACAACATA AGACAATATA ATCATGTGCT TTTAACATAT TTGAGAATAA AAAGGACTAG  1140
CC
```

Seq ID NO: 24 Protein sequence
Protein Accession #: NP_006519

```
1          11         21         31         41         51
|          |          |          |          |          |
MDPARPLGLS ILLLFLTEAA LGDAAQEPTG NNAEICLLPL DYGPCRALLL RYYYDRYTQS    60
CRQFLYGGCE GNANNFYTWE ACDDACWRIE KVPKVCRLQV SVDDQCEGST EKYFFNLSSM   120
TCEKFFSGGC HRNRIENRFP DEATCMGFCA PKKIPSFCYS PKDEGLCSAN VTRYYFNPRY   180
RTCDAFTYTG CGGNDNNFVS REDCKRACAK ALKKKKKMPK LRFASRIRKI RKKQF
```

Seq ID NO: 25 DNA sequence
Nucleic Acid Accession #: NM_005458.1
Coding sequence: 1..2825

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGCTTCCC CGCGGGAGGTC CGGGCAGCCA GGGCGGCCGC CGCCGCCGCC ACCGCCGCCC    60
GCGCGCCTGC TACTGCTACT GCTGCTGCCG CTGCTGCTGG CCCAGCAGCC CTCTGGCGCC CGGGGCCTGG   120
GGCTGGGCGC GGGGGCGCCC CCGGCCGCCG CCCAGCAGCC CGCCGCTCTC CATCATGGGC   180
CTCATGCCGC TCACCAAGGA GGTGGCCAAG GGCAGCATCG GGCGCGGTGT GCTCCCCGCC   240
GTGGAACTGG CCATCGAGCA GATCCGCAAC GAGTCACTCC TGCGCCCCTA CTTCCTCGAC   300
CTGCGGCTCT ATGACACGGA GTGCGACAAC GCAAAAGGGT TGAAAGCCTT CTACGATGCA   360
ATAAAATACG GGCCGAACCA CTTGATGGTG TTTGGAGGCG TCTGTCCATC CGTCACATCC   420
ATCATTGCAG AGTCCCTCCA AGGCTGGAAT CTGGTGCAGC TTTCTTTTGC TGCAACCACG   480
CCTGTTCTAG CCGATAAGAA AAAATACCCT TATTTCTTTC GGACCGTCCC ATCAGACAAT   540
GCGGTGAATC CAGCCATTCT GAAGTTGCTC AAGCACTACC AGTGGAAGCC CGTGGGCACG   600
CTGACGCAAG ACGTTCAGAG GTTCTCTGAG GTGCGGAATG ACCTGACTGG AGTTCTGTAT   660
GGGCGAGGACA TTGAGATTTC AGACACCGAG AGCTTCTCCA ACGATCCCTG TACCAGTGTC   720
AAAAAGCTGA AGGGGAATGA TGTGCGGATC ATCCTTGGCC AGTTTGACCA GAATATGGCA   780
GCAAAAGTGT TCTGTTGTGC ATACGAGGAG AACATGTATG GTAGTAAATA TCAGTGGATC   840
ATTCCGAGGC GGTACGAGCC TTCTTGGTGG GAGCAGGTGC ACAGGGAAGC CAACTCATCC   900
CGCTGCCTCC GGAAGAATCG GCTTGCTGCC ATGGAGGGCT ACATTGGCGT GGAATTCGAG   960
CCCCTGAGCT CCAAGCAGAT CAAGACCATC TCAGGAAAGA CTCCACAGCA GTATGAGAGA  1020
GAGTACAACA ACAAGCGGTC AGGCGTGGGG CCCAGCAAGT TCCACGGGTA CGCCTACGAT  1080
GGCATCTGGG TCATCGCCAA GACACTGCAG AGGGCCATGG AGACACTGCA TGCCAGCAGC  1140
CGGCACCAGC GGATCCAGGA CTTCAACTAC ACGGACCACA CGCTGGGCAG GATCATCCTC  1200
AATGCCATGA ACGAGACCCA CTTCTTCGGG GTCACGGGTC AAGTTGTATT CCGGAATGGG  1260
GAGAGAATGG GGACCATTAA ATTTACTCAA TTTCAAGACA GCAGGGAGGT GAAGGTGGGA  1320
GAGTACAACG CTGTGGCCGA CACACTGGAG ATCATCAATG ACACCATCAG GTTCCAAGGA  1380
TCCGAACCAC CAAAAGACA GACCATCATC CTGGAGCAGC TGCGGAAGAT CTCCCTACCT  1440
CTCTACACGA TCCTCTCTGC CCTCACCATC CTCGGGATGA TCATGGCCAG TGCTTTTCTC  1500
TTCTTCAACA TCAAGAACCG GAATCAGAAG CTCATAAAGA TGTCGAGTCC ATACATGAAC  1560
AACCTTATCA TCCTTGGAGG GATGCTCTCC TATGCTTCCA TATTTCTCTT TGGCCTTGAT  1620
GGATCCTTTG TCTCTGAAAA GACCTTTGAA ACACTTTGCA ACCGTGGGCT ACACGACCGC  1680
ACCGTGGGCT ACACGACCGC TTTTGGGGCC ATGTTTGCAA AGACCTGGAG AGTCCACGCC  1740
ATCTTCAAAA ATGTGAAAAT GAAGAAGAAG ATCATCAAGG ACCAGAAACT GCTTGTGATC  1800
GTGGGGGGCA TGCTGCTGAT CGACCTGTGT ATCCTGATCT GCTGGCAGGC TGTGGACCCC  1860
CTGCGAAGGA CAGTGGAGAA GTACAGCATG GAGCCGGACC CAGCAGGACG GGATATCTCC  1920
ATCCGCCCTC TCCTGGAGCA CTGTGAGAAC ACCCATATGA CCATCTGGCT TGGCATCGTC  1980
TATGCCTACA AGGGACTTCT CATGTTGTTC GGTTGTTTCT TAGCTTGGGA GACCCGCAAC  2040
GTCAGCATCC CCGCACTCAA CGACAGCAAG TACATCGGGA TGAGTGTCTA CAAGTGGGG  2100
ATCATGTGCA TCATCGGGGC CGCTGTCTCC TTCCTGACCC GGGACCAGCC CAATGTGCAG  2160
TTCTGCATCG TGGCTCTGGT CATCATCTTC TGCAGCACCA TCACCCTCTG CCTGGTATTC  2220
GTGCCGAAGC TCATCACCCT GAGAACAAAC CCAGATGCAG CAACGCAGAA CAGGCGGATTC  2280
CAGTTCACTC AGAATCAGAA GAAAGAAGAT TCTAAAACGT CCACCTCGGT CACCAGTGTG  2340
AACCAAGCCA GCACATCCCG CCTGGAGGGC CTACAGTCAG AAAACCATCG CCTGCGAATG  2400
```

957

```
AAGATCACAG AGCTGGATAA AGACTTGGAA GAGGTCACCA TGCAGCTGCA GGACACACCA   2460
GAAAAGACCA CCTACATTAA ACAGAACCAC TACCAAGAGC TCAATGACAT CCTCAACCTG   2520
GGAAACTTCA CTGAGAGCAC AGATGGAGGA AAGGCCATTT TAAAAAATCA CCTCGATCAA   2580
AATCCCCAGC TACAGTGGAA CACAACAGAG CCCTCTCGAA CATGCAAAGA TCCTATAGAA   2640
GATATAAACT CTCCAGAACA CATCCAGCGT CGGCTGTCCC TCCAGCTCCC CATCCTCCAC   2700
CACGCCTACC TCCCATCCAT CGGAGGCGTG GACGCCAGCT GTGTCAGCCC CTGCGTCAGC   2760
CCCACCGCCA GCCCCCGCCA CAGACATGTG CCACCCTCCT TCCGAGTCAT GGTCTCGGGC   2820
CTGTAA
```

Seq ID NO: 26 <u>Protein sequence</u>
Protein Accession #: NP_005449.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MASPRRSGQP GRPPPPPPPP ARLLLLLLLP LLLPLAPGAW GNARGAPRPP PSSPPLSIMG   60
LMPLTKEVAK GSIGRGVLPA VELAIEQIRN ESLLRPYPLD LRLYDTECDN AKGLKAFYDA   120
IKYGPNHLMV FGGVCPSVTS IIAESLQGWN LVQLSFAATT PVLADKKKYP YFFRTVPSDN   180
AVNPAILKLL KHYQWKRVGT LTQDVQRFSE VRNDLTGVLY GEDIEISDTE SFSNDPCTSV   240
KKLKGNDVRI ILGQFDQNMA AKVFCCAYEE NMYGSKYQWI IPGWYEPSWW EQVHTEANSS   300
RCLRKNLLAA MEGYIGVDFE PLSSKQIKTI SGKTPQQYER EYNNKRSGVG PSKFHGYAYD   360
GIWVIAKTLQ RAMETLHASS RHQRIQDFNY TDHTLGRIIL NAMNETNFFG VTGQVVFRNG   420
ERMGTIRFTQ FQDSREVKVG EYNAVADTLE IINDTIRFQG SEPPKDKTII LEQLRKISLP   480
LYSILSALTI LGMIMASAPL FFNIKNRNQK LIKMSSPYMN NLIILGGMLS YASIFLFGLD   540
GSFVSEKTFE TLCTVRTWIL TVGYTTAFGA MFAKTWRVHA IFKNVKMKKK IIKDQKLLVI   600
VGGMLLIDLC ILICWQAVDP LRRTVEKYSM SPDPAGRDIS IRPLLEHCEN THMTIWLGIV   660
YAYKGLLMLF GCFLAWETRN VSIPALNDSK YIGMSVYNVG IMCIIGAAVS FLTRDQPNVQ   720
FCIVALVIIF CSTITLCLVF VPKLITLRTN PDAATQNRRP QPTQNQKKED SKTSTSVTSV   780
NQASTSRLEG LQSENHRLRM XITELDKDLE EVTMQLQDTP EKTTYIKQNH YQELNDILNL   840
GNFTESTDGG KAILKNHLDQ NPQLQWNTTE PSRTCKDPIE DINSPEHIQR RLSLQLPILH   900
HAYLPSIGGV DASCVSPCVS PTASPRHRHV PPSPRVMVSG L                      901
```

Seq ID NO: 27 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_000450.1
Coding sequence: 117..1949

```
1          11         21         31         41         51
|          |          |          |          |          |
CCTGAGACAG AGGCAGCAGT GATACCCACC TGAGAGATCC TGTGTTTGAA CAACTGCTTC   60
CCAAAACGGA AAGTATTTCA AGCCTAAACC TTTGGGTGAA AAGAACTCTT GAAGTCATGA   120
TTGCTTCACA GTTTCTCTCA GCTCTCACTT TGGTGAGCTT CATTAAAGAG AGTGGAGCCT   180
GGTCTTACAA CACCTCCACG GAAGCTATGA CTTATGATGA GGCCAGTGCT TATTGTCAGC   240
AAAGGTACAC ACACCTGGTT GCAATTCAAA ACAAAGAAGA GATTGAGTAC CTAAACTCCA   300
TATTGAGCTA TTCACCAAGT TATTACTGGA TTGGAATCAG AAAAGTCAAC AATGTGTGGG   360
TCTGGGTAGG AAACCCAGAA CCTCTGACAG AAGAAGCCAA GAACTGGGCT CCAGGTGAAC   420
CCAACAATAG GCAAAAAGAT GAGGGACTGCG TGGAGATCTA CATCAAGAGA GAAAAAGATG   480
TGGGCATGTG GAATGATGAG AGGTGCAGCA AGAAGAAGCT TGCCCTATGC TACACAGCTG   540
CCTGTACCAA TACATCCTGC AGTGGCCACG GTGAATGTGT AGAGACCATC AATAATTACA   600
CTTGCAAGTG TGACCCTGGC TTCAGTGGAC TCAAGTGTGA GCAAATTGTG AACTGTACAG   660
CCCTGGAATC CCCTGAGCAT GGAAGCCTGG TTTGCAGTCA CCCACTGGGA AACTTCAGCT   720
ACAATTCTTC CTGCTCTATC AGCTGTGATA GGGGTTACCT GCCAAGCAGC ATGGAGACCA   780
TGCAGTGTAT GTCCTCTGGA GAATGGAGTG CTCCTATTCC AGCCTGCAAT GTGGTTGAGT   840
GTGATGCTGT GACAAATCCA GCCAATGGGT TCGTGGAATG TTTCCAAAAC CCTGGAAGCT   900
TCCCATGGAA CACAACCTGT ACATTTGACT GTGAAGAAGG ATTTGAACTA ATGGGAGCCC   960
AGAGCCTTCA GTGTACCTCA TCTGGGAATT GGGACAACGA GAAGCCAACG TGTAAAGCTG   1020
TGACATGCAG GGCCGTCCGC CAGCCTCAGA ATGGCTCTGT GAGGTGCAGC CATTCCCCTG   1080
CTGGGACAGC CACCTTCAAA TCATCCTGCA ACTTCACCTG TGAGGAAGGC TTCATGTTGC   1140
AGGGACCAGC CCAGGTTGAA TGCACCACTC AAGGGCAGTG GACACAGCAA ATCCCAGTTT   1200
GTGAAGCTTT CCAGTGCACA GCCTTGTCCA ACCCCGAGCG AGGCTACATG AATTGTCTTC   1260
CTAGTGTCTC TGGCAGTTTC CGTTATGGGT CCAGCTGTGA GTTCTCCTGT GAGCAGGGTT   1320
TTGTGTTGAA GGGATCCAAA AGGCTCCAAT GTGGCCCCAC AGGGGAGTGG GACAACGAGA   1380
AGCCCACATG TGAAGCTGTG AGATGCGATG CTGTCCACCA GCCCCCGAAG GGTTTGGTGA   1440
GGTGTGCTCA TTCCCCTATT GGAGAATTCA CCTACAAGTC CTCTTGTGCC TTCAGCTGTG   1500
AGGAGGGGATT TGAATTATAT GGGATCAACTC AACTTGAGTG CACATCTCAG GGACAATGGA   1560
CAGAAGAGGT TCCTTCCTGC CAAGTGGTAA AATGTTCAAG CCTGGCAGTT CCGGGAAAGA   1620
TCAACATGAG CTGCAGTGGG GAGCCCGTGT TTGGCACTGT GTGCAAGTTC GCCTGTCCTG   1680
AAGGATGGAC GCTCAATGGC TCTGCAGCTC GGACATGTGG AGCCACAGGA CACTGGTCTG   1740
GCCTGCTACC TACCTGTGAA GCTCCCACTG AGTCCAACAT TCCCTTGGTA GCTGGACTTT   1800
CTGCTGCTGG ACTCTCCCTC CTGACATTAG CACCATTTCT CCTCTGGCTT CGGAAATGCT   1860
TACGGAAAGC AAAGAAATTT GTTCCTGCCA GCAGCTGCCA AAGCCTTGAA TCAGACGGAA   1920
GCTACCAAAA GCCTTCTTAC ATCCTTTAAG TTCAAAAGAA TCAGAAACAG GTGCATCTGG   1980
GGAACTAGAG GGATACACTG AAGTTAACAG AGACAGATAA CTCTCCTCGG GTCTCTGGCC   2040
CTTCTTGCCT ACTATGCCAG ATGCCTTTAT GGCTGAAACC CGAACCACTTC 2100
AATAGATCAA AGTCCAGCAG GCAAGGACGG CCTTCAACTG AAAAGACTCA GTGTTCCCTT   2160
TCCTACTCTC AGGATCAAGA AAGTGTTGGC TAATGAAGGG AAAGGATATT TTCTTCCAAG   2220
CAAAGGTGAA GAGACCAAGA CTCTGAAATC TCAGAATTCC TTTTCTAACT CTCCCTTGCT   2280
CGCTGTAAAA TCTTGGCACA GAAACACAAT ATTTTGTGGC TTTCTTTCTT TTGCCCTTCA   2340
CAGTGTTTCG ACAGCTGATT ACACAGTTGC TGTCATAAGA ATGAATAATA ATTATCCAGA   2400
GTTTAGAGGA AAAAAATGAC TAAAAAATAT ATAACTTAAA AAAATGACAG ATGTTGAATG   2460
CCCACAGGCA AATGCATGGA GGGTTGTTAA TGGTGCAAAT CCTACTGAAT GCTCTGTGCG   2520
AGGGTTACTA TGCACAATTT AATCACTTTC ATCCCTATGG GATTCAGTGC TTCTTAAAGA   2580
GTTCTTAAGG ATTGTGATAT TTTTACTTGC ATTGAATATA TTATAATCTT CCATACTTCT   2640
TCATTCAATA CAAGTGTGGT AGGGACTTAA AAAACTTGTA AATGCTGTCA ACTATGATAT   2700
GGTAAAAGTT ACTTATTCTA GATTACCCCC TCATTGTTTA TTAACAAATT ATGTTACATC   2760
TGTTTTAAAT TTATTTCAAA AAGGGAAACT ATTGTCCCCT AGCAAGGCAT GATGTTAACC   2820
```

```
          AGAATAAAGT TCTGAGTGTT TTTACTACAG TTGTTTTTTG AAAACATGGT AGAATTGGAG   2880
          AGTAAAAACT GAATGGAAGG TTTGTATATT GTCAGATATT TTTTCAGAAA TATGTGGTTT   2940
          CCACGATGAA AAACTTCCAT GAGGCCAAAC GTTTTGAACT AATAAAAGCA TAAATGCAAA   3000
          CACACAAAGG TATAATTTTA TGAATGTCTT TGTTGGAAAA GAATACAGAA AGATGGATGT   3060
5         GCTTTGCATT CCTACAAAGA TGTTTGTCAG ATGTGATATG TAAACATAAT TCTTGTATAT   3120
          TATGGAAGAT TTTAAATTCA CAATAGAAAC TCACCATGTA AAAGAGTCAT CTGGTAGATT   3180
          TTTAACGAAT GAAGATGTCT AATAGTTATT CCCTATTTGT TTTCTTCTGT ATGTTAGGGT   3240
          GCTCTGGAAG AGAGGAATGC CTGTGTGAGC AAGCATTTAT GTTTATTTAT AAGCAGATTT   3300
          AACAATTCCA AAGGAATCTC CAGTTTTCAG TTGATCCATG GCAATGAAAA ATTCTCAGTC   3360
10        AGTAATTGCC AAAGCTGCTC TAGCCTTGAG GAGTGTGAGA ATCAAAACTC TCCTACACTT   3420
          CCATTAACTT AGCATGTGTT GAAAAAAAAA GTTTCAGAGA AGTTCTGGCT GAACACTGGC   3480
          AACGACAAAG CCAACAGTCA AAACAGAGAT GTGATAAGGA TCAGAACAGC AGAGGTTCTT   3540
          TTAAAGGGGC AGAAAAACTC TGGGAAATAA GAGAGAACAA CTACTGTGAT CAGGCTATGT   3600
          ATGGAATACA GTGTTATTTT CTTTGAAATT GTTTAAGTGT TGTAAATATT TATGTAAACT   3660
15        GCATTAGAAA TTAGCTGTGT GAAATACCAG TGTGGTTTGT GTTTGAGTTT TATTGAAGAT   3720
          TTTAAATTAT AACTTAAAAT ATTTTATAAT TTTTAAAGTA TATATTTATT TAAGCTTATG   3780
          TCAGACCTAT TTGACATAAC ACTATAAAGG TTGACAATAA ATGTGCTTAT GTTT
```

20 Seq ID NO: 28 <u>Protein sequence</u>
Protein Accession #: NP_000441.1

```
          1          11         21         31         41         51
          |          |          |          |          |          |
25        MIASQFLSAL TLVLLIKESG AWSYNTSTEA MTYDEASAYC QQRYTHLVAI QNKEEIEYLN   60
          SILSYSPSYY WIGIRKVNNV WVWVGTQKPL TEEAKNWAPG EPNNRQKDED CVEIYIKREK   120
          DVGMWNDERC SKKKLALCYT AACTNTSCSG HGECVETINN YTCKCDPGFS GLKCEQIVNC   180
          TALESPEHGS LVCSHPLGNF SYNSSCSISC DRGYLPSSME TMQCMSSGEW SAPIPACNVV   240
          ECDAVINPAN GFVECPQNPG SPPWNTTCTF DCEEGFELMG AQSLQCTSSG NWDNEKPTCK   300
30        AVTCRAVRQP QNGSVRCSHS PAGEPTFKSS CNFTCEEGFM LQGPAQVECT TQGQWTQQIP   360
          VCEAFQCTAL SNPERGYMNC LPSASGSFRY GSSCEFSCEQ GFVLRGSKRL QCGPTGEWDN   420
          EKPTCEAVRC DAVHQPPKGL VRCAHSPIGE FTYKSSCAFS CEEGFELYGS TQLECTSGGQ   480
          WTEEVPSCQV VKCSSLAVPG KINMSCSGEP VFGTVCKFAC PEGWTLNGSA ARTCGATGHW   540
          SGLLPTCEAP TESNIPLVAG LSAAGLSLLT LAPFLLWLRK CLRKAKKFVP ASSCQSLESD   600
35        GSYQKPSYIL
```

Seq ID NO: 29 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_007036
Coding sequence: 56..610

```
40        1          11         21         31         41         51
          |          |          |          |          |          |
          CTTCCCACCA GCAAAGACCA CGACTGGAGA GCCGAGCCGG AGGCAGCTGG GAAACATGAA   60
          GAGCGTCTTG CTGCTGACCA CGCTCCTCGT GCCTGCCACAC CTGGTGGCCG CCTGGAGCAA   120
          TAATTATGCG GTGGACTGCC CTCAACACTG TGACAGCAGT GAGTGCAAAA GCAGCCCGCG   180
45        CTGCAAGAGG ACAGTGCTCG ACGACTGTGG CTGCTGCCGA GTGTGCGCTG CAGGGCGGGG   240
          AGAAACTTGC TACCGCACAG TCTCAGGCAT GGATGGCATG AAGTGTGGCC CGGGGCTGAG   300
          GTGTCAGCCT TCTAATGGGG AGGATCCTTT TGGTGAAGGA TTTGGTATCT GCAAAGACTG   360
          TCCCTACGGC ACCTTCGGGA TGGATTGCAG AGAGACCTGC AACTGCCAGT CAGGGCATCG   420
          TGACAGGGGG ACGGGAAAAT GCCTGAAATT CCCCTTCTTC CAATATTCAG TAACCAAGTC   480
50        TTCCAACAGA TTTGTTTCTC TCACGGAGCA TGACATGGCA TCTGGAGATG GCAATATTGT   540
          GAGAGAAGAA GTTGTGAAAG AGAATGCTGC CGGGTCTCCC GTAATGAGGA AATGGTTAAA   600
          TCCACGCTGA TCCCGGCTGT GATTTCTGAG AGAAGGCTCT ATTTTCGTGA TTGTTCAACA   660
          CACAGCCAAC ATTTTAGGAA CTTTCTAGAT ATAGCATAAG TACATGTAAT TTTTGAAGAT   720
          CCAAATTGTG ATGCATGGTG GATCCAGAAA ACAAAAAGTA GGATACTTAC AATCCATAAC   780
55        ATCCATATGA CTGAACACTT GTATGTGTTT GTTAAATATT CGAATGCATG TAGATTTGTT   840
          AAATGTGTGT GTATAGTAAC ACTGAAGAAC TAAAAATGCA ATTTAGGTAA TCTTACATGG   900
          AGACAGGTCA ACCAAAGAGG GAGCTAGGCA AAGCTGAAGA CCGCAGTGAG TCAAATTAGT   960
          TCTTTGACTT TGATGTACAT TAATGTTGGG ATATGGAATG AAGACTTAAG AGCAGGAGAA   1020
          GATGGGGAGG GGGTGGGAGT GGGAAATAAA ATATTTAGCC CTTCCTTGGT AGGTAGCTTC   1080
60        TCTAGAATTT AATTGTGCTT TTTTTTTTTT TTTGGCTTTG GGAAAAGTCA AAATAAAACA   1140
          ACCAGAAAAC CCCTGAAGGA AGTAAGATGT TTGAAGCTTA TGGAAATTTG AGTAACAAAC   1200
          AGCTTTGAAC TGAGAGCAAT TTCAAAAGGC TGCTGATGTA GTTCCCGGGT TACCTGTATC   1260
          TGAAGGACGG TTCTGGGGCA TAGGAAACAC ATACACTTCC ATAAATAGCT TTAACGTATG   1320
          CCACCTCAGA GATAAATCTA AGAAGTATTT TACCCACTGG TGGTTTGTGT GTGTATGAAG   1380
65        GTAAATATTT ATATATTTTT ATAAATAAAT GTGTTAGTGC AAGTCATCTT CCCTACCCAT   1440
          ATTTATCATC CTCTTGAGGA AAGAAATCTA GTATTATTTG TTGAAAATGG TTAGAATAAA   1500
          AACCTATGAC TCTATAAGGT TTTCAAACAT CTGAGGCATG ATAAATTTAT TATCCATAAT   1560
          TATAGGAGTC ACTCTGGATT TCAAAAAATG TCAAAAAATG AGCAACAGAG GGACCTTATT   1620
          TAAACATAAA TGCTGTGACT TCGGTGAATT TTCAATTTAA GGTATGAAAA TAAGTTTTTA   1680
70        GGAGGTTTGT AAAAGAAGAA TCAATTTTCA GCAGAAAACA TGTCAACTTT AAAAATATAGG   1740
          TGGAATTAGG AGTATATTTG AAAGAATCTT AGCACAAACA GGACTGTTGT ACTAGATGTT   1800
          CTTAGGAAAT ATCTCAGAAG TATTTTATTT GAAGTGAAGA ACTTATTTAA GAATTATTTC   1860
          AGTATTTACC TGTATTTTAT TCTTGAAGTT GGCCAACAGA GTTGTGAATG TGTGTGGAAG   1920
          GCCTTTGAAT GTAAAGCTGC ATAAGCTGTT AGGTTTTGTT TTAAAAGGAC ATGTTTATTA   1980
75        TTGTTCAATA AAAAAGAACA AGATAC
```

Seq ID NO: 30 <u>Protein sequence</u>
Protein Accession #: NP_008967.1

```
80        1          11         21         31         41         51
          |          |          |          |          |          |
          MKSVLLLTTL LVPAHLVAAW SNNYAVDCPQ HCDSSECKSS PRCKRTVLDD CGCCRVCAAG   60
          RGETCYRTVS GMDGMKCGPG LRCQPSNGED PFGEEFGICK DCPYGTFGMD CRETCNCQSG   120
          ICDRGTGKCL KFPPFQYSVT KSSNRFVSLT EHDMASGDGN IVREEVVKEN AAGSFVMRKN   180
```

959

LNPR

Seq ID NO: 31 DNA sequence
Nucleic Acid Accession #: NM_000963
Coding sequence: 135..1949

```
1          11         21         31         41         51
|          |          |          |          |          |
CAATTGTCAT ACGACTTGCA GTGAGCGTCA GGAGCACGTC CAGGAACTCC TCAGCAGCGC    60
CTCCTTCAGC TCCACAGCCA GACGCCCTCA GACAGCAAAG CCTACCCCCG CGCCGCGCCC   120
TGCCCGCCGC TCGGATGCTC GCCCGCGCCC TGCTGCTGTG CGCGGGTCCTG GCGCTCAGCC   180
ATACAGCAAA TCCTTGCTGT TCCCACCCAT GTCAAAACCG AGGTGTATGT ATGAGTGTGG   240
GATTTGACCA GTATAAGTGC GATTGTACCC GGACAGGATT CTATGGAGAA AACTGCTCAA   300
CACCGGAATT TTTGACAAGA ATAAAATTAT TTCTGAAACC CACTCCAAAC ACAGTGCACT   360
ACATACTTAC CCACTTCAAG GGATTTTTGGA ACGTTGTGAA TAACATTCCC TTCCTTCGAA   420
ATGCAATTAT GAGTTATGTC TTGACATCCA GATCACATTT GATTGACAGT CCACCAACTT   480
ACAATGCTGA CTATGGCTAC AAAAGCTGGG AAGCCTTCTC TAACCTCTCC TATTATACTA   540
GAGCCCTTCC TCCTGTGCCT GATGATTGCC CGACTCCCTT GGGTGTCAAA GGTAAAAAGC   600
AGCTTCCTGA TTCAAATGAG ATTGTGGAAA AATTGCTTCT AAGAAGAAAG TTCATCCCTG   660
ATCCCCAGGG CTCAAACATG ATGTTTGCAT TCTTTGCCCA GCACTTCACG CATCAGTTTT   720
TCAAGACAGA TCATAAGCGA GGGCCAGCTT TCACCAACGG GCTGGGCCAT GGGGTGGACT   780
TAAATCATAT TTACGGTGAA ACTCTGGCTA GACAGCGTAA ACTGCGCCTT TTCAAGGATG   840
GAAAAATGAA ATATCAGATA ATTGATGGAG AGATGTATCC TCCCACAGTC AAAGATACTC   900
AGGCAGAGAT GATCTACCCT CCTCAAGTCC CTGAGCATCT ACGGTTTGCT GTGGGGCAGG   960
AGGTCTTTGG TCTGGTGCCT GGTCTGATGA TGTATGCCAC AATCTGGCTG CGGGAACACA  1020
ACAGAGTATG CGATGTGCTT AAACAGGAGC ATCCTGAATG GGGTGATGAG CAGTTGTTCC  1080
AGACAAGCAG GCTAATACTG ATAGGAGAGA CTATTAAGAT TGTGATTGAA GATTATGTGC  1140
AACACTTGAG TGGCTATCAC TTCAAACTGA AATTTGACCC AGAACTACTT TTCAACAAAC  1200
AATTCCAGTA CCAAAATCGT ATTGCTGCTG AATTAACAC CCTCTATCAC TGGCATCCCC  1260
TTCTGCCTGA CACCTTTCAA ATTCATGACC AGAAATACAA CTATCAACAG TTTATCTACA  1320
ACAACTCTAT ATTGCTGGAA CATGGAATTA CCCAGTTTGT TGAATCATTC ACCAGGCAAA  1380
TTGCTGGCAG GGTTGCTGGT GGTAGGAATG TTCCACCCGC AGTACAGAAA GTATCACAGG  1440
CTTCCATTGA CCAGAGCAGG CAGATGAAAT ACCAGTCTTT TAATGAGTAC CGCAAACGCT  1500
TTATGCTGAA GCCCTATGAA TCATTTGAAG AACTTACAGG AGAAAAGGAA ATGTCTGCAG  1560
AGTTGGAAGC ACTCTATGGT GACATCGATG CTGTGGAGCT CTTCTGGTAG 1620
AAAAGCCTCG GCCAGATGCC ATCTTTGGTG AAACCATGGT AGAAGTTGGA GCACCATTCT  1680
CCTTGAAAGG ACTTATGGGT AATGTTATAT GTTCTCCTGC CTACTGGAAG CCAAGCACTT  1740
TTGGTGGAGA AGTGGGTTTT CAAATCATCA ACACTGCCTA AATTCAGTCT CTCATCTGCA  1800
ATAACGTGAA GGGCTGTCCC TTTACTTCAT TCAGTGTTCC AGATCCAGAG CTCATTAAAA  1860
CAGTCACCAT CAATGCAAGT TCTTCCCGCT CCGGACTAGA TGATATCAAT CCCACAGTAC  1920
TACTAAAAGA ACGTTCGACT GAACTGTAGA AGTCTAATGA TCATATTTAT TTATTTATAT  1980
GAACCATGTC TATTAATTTA ATTATTTAAT AATATTTATA TTAAACTCCT TATGTTACTT  2040
AACATCTTCT GTAACAGAAG TCAGTACTCC TGTTGCGGAG AAAGGAGTCA TACTTGTGAA  2100
GACTTTTATG TCACTACTCT AAAGATTTTG CTGTTGCTGT TAAGTTTGGA AAACAGTTTT  2160
TATTCTGTTT TATAAACCAG AGAGAAATGA GTTTTGACGT CTTTTTACTT GAATTTCAAC  2220
TTATATTATA AGAACGAAAG TAAAGATGTT TGAATACTTA AACACTATCA CAAGATGGCA  2280
AAATGCTGAA AGTTTTTACA CTGTCGATGT TTCCAATGCA TCTTCCATGA TGCATTAGAA  2340
GTAACTAATG TTTGAAATTT TAAAGTACTT TTGGTTATTT TTCTGTCATC AAACAAAAAC  2400
AGGTATCAGT GCATTATTAA ATGAATATTT AAATTAGACA TTACCAGTAA TTTCATGTCT  2460
ACTTTTTAAA ATCAGCAATG AAACAATAAT TTGAAATTTC TAAATTCATA GGGTAGAATC  2520
ACCTGTAAAA GCTTGTTTGA TTTCTTAAAG TTATTAAACT TGTACATATA CCAAAAAGAA  2580
GCTGTCTTGG ATTTAAATCT GTAAAATCAG ATGAAATTTT ACTACAATTG CTTGTTAAAA . 2640
TATTTTATAA GTGATGTTCC TTTTTCACCA AGAGTATAAA CCTTTTTAGT GTGACTGTTA  2700
AAACTTCCTT TTAAATCAAA ATGCCAAATT TATTAAGGTG GTGGAGCCAC TGCAGTGTTA  2760
TCTCAAAATA AGAATATTTT GTTGAGATAT TCCAGAATTT GTTTATATGG CTGGTAACAT  2820
GTAAATCTA TATCAGCAAA AGGGTCTACC TTTAAAATAA GCAATAACAA AGAAGAAAC  2880
CAAATTATTG TTCAAATTTA GGTTTAAACT TTTGAAGCAA ACTTTTTTTT ATCCTTGTGC  2940
ACTGCAGGCC TGGTACTCAG ATTTTGCTAT GAGGTTAATG AAGTACCAAG CTGTGCTTGA  3000
ATAACGATAT GTTTTCTCAG ATTTTCTGTT GTACAGTTTA ATTTAGCAGT CCATATCACA  3060
TTGCAAAAGT AGCAATGACC TCATAAAATA CCTCTTCAAA TCGTTAAAT TCATTTCACA  3120
CATTAATTTT ATCTCAGTCT TGAAGCCAAT TCAGTAGGTG CATTGGAATC AAGCCTGGCT  3180
ACCTGCATGC TGTTCCTTTT CTTTTCTTCT TTTAGCCATT TTGCTAAGAG ACACAGTCTT  3240
CTCATCACTT CGTTTCTCCT ATTTTGTTTT ACTAGTTTTA AGATCAGAGT TCACTTTCTT  3300
TGGACTCTGC CTATATTTTC TTACCTGAAC TTTTGCAAGT TTTCAGGTAA ACCTCAGCTC  3360
AGGACTGCTA TTTAGCTCCT CTTAAGAAGA TTAAAGAGA AAAAAAAGG CCCTTTTAAA  3420
AATAGTATAC ACTTATTTTA AGTGAAAAGC AGAGAATTTT ATTTATAGCT AATTTTAGCT  3480
ATCTGTAACC AAGATGGATG CAAAGAGGCT AGTGCCTCAG AGAGAACTGT ACGGGGTTTG  3540
TGACTGGAAA AAGTTACGTT CCCATTCTAA TTAATGCCCT TTCTTATTTA AAAACAAAAC  3600
CAAATGATAT CTAAGTAGTT CTCAGCAATA ATAATAATGA CGATAATACT TCTTTTCCAC  3660
ATCTCATTGT CACTGACATT TAATGGTACT GTATATTATT TAATTAATTG AAGATTATTA  3720
TTTATGTCTT ATTAGGACAC TATGGTTATA AACTGTGTTT AAGCCTACAA TCATTGATTT  3780
TTTTTTGTTA TGTCACAAТC AGTATATTTT CTTTGGGGTT ACCTCTCTGA ATATTATGTA  3840
AACAATCCAA AGAAATGATT GTATTAAGAT TTGTGAATAA ATTTTTAGAA ATCTGATTGG  3900
CATATTGAGA TATTTAAGGT TGAATGTTTG TCCTTAGGAT AGGCCTATGT GCTAGCCCAC  3960
AAAGAATATT GTCTCATTAG CCTGAATGTG CCATAAGACT GACCTTTTAA AATGTTTTGA  4020
GGGATCTGTG GATGCTTCGT TAATTTGTTC AGCCACAATT TATTGAGAAA ATATTCTGTG  4080
TCAAGCACTG TGGGTTTTAA TATTTTTAAA TCAAACGCTG ATTACAGATA ATAGTATTTA  4140
TATAAATAAT TGAAAAAAAT TTTCTTTTGG GAAGAGGGAG AAAATGAAAT AAATATCATT  4200
AAAGATAACT CAGGAGAATC TTCTTTACAA TTTTACGTTT AGAATGTTTA AGGTTAAGAA  4260
AGAAATAGTC AATATGCTTG TATAAAAACAC TGTTCACTGT TTTTTTTAAA AAAAAACTTT  4320
GATTTGTTAT TAACATTGAT CTGCTGACAA AACCTGGGAA TTTGGGTTGT GTATGCGAAT  4380
GTTTCAGTGC CTCAGACAAA TGTGTATTTA ACTTATGTAA AAGATAAGTC TGGAAATAAA  4440
TGTCTGTTTA TTTTTGTACT ATTTA
```

Seq ID NO: 32 Protein sequence
Protein Accession #: NP_000954

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     MLARALLLCA VLALSHTANP CCSHPCQNRG VCMSVGFDQY KCDCTRTGFY GENCSTPEFL   60
     TRIKLFLKPT PNTVHYILTH FKGFWNVVNN IPFLRNAIMS YVLTSRSHLI DSPPTYNADY  120
     GYKSWEAFSN LSYYTRALPP VPDDCPTPLG VKGKKQLPDS NEIVEKLLLR RKFIPDPQGS  180
     NMMFAFFAQH FTHQFFKTDH KRGPAPTNGL GHGVDLNHIY GETLARQRKL RLFKDGKMKY  240
     QIIDGEMYPP TVKDTQAEMI YPPQVPEHLR FAVGQEVFGL VPGLMMYATI WLREHNRVCD  300
     VLKQEHPEWG DEQLFQTSRL ILIGETIKIV IEDYVQHLSG YHFKLKFDPE LLFNKQFQYQ  360
     NRIAAEFNTL YHWHPLLPDT FQIHDQKYNY QQFIYNNSIL LEHGITQFVE SFTRQIAGRV  420
     AGGRNVPPAV QKVSQASIDQ SRQMKYQSFN EYRKRFMLKP YESFEELTGE KEMSAELEAL  480
     YGDIDAVELY PALLVEKPRP DAIFGETMVE VGAPFSLKGL MGNVICSPAY WKPSTFGGEV  540
     GFQIINTASI QSLICNNVKG CPFTSFSVPD PELIKTVTIN ASSSRSGLDD INPTVLLKER  600
     STEL
```

Seq ID NO: 33 DNA sequence
Nucleic Acid Accession #: NM_001508.1
Coding sequence: 1..1361

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     ATGGCTTCAC CCAGCCTCCC GGGCAGTGAC TGCTCCCAAA TCATTGATCA CAGTCATGTC   60
     CCCGAGTTTG AGGTGGCCAC CTGGATCAAA ATCACCCTTA TTCTGGTGTA CCTGATCATC  120
     TTCGTGATGG GCCTTCTGGG GAACAGCGTC ACCATTCGGG TCACCCAGGT GCTGCAGAAG  180
     AAAGGATACT TGCAGAAGGA GGTGACAGAC CACATGGTGA GTTTGGCTTG CTCGGACATC  240
     TTGGTGTTCC TCATCGGCAT GCCCATGGAG TTCTACAGCA TCATCTGGAA TCCCCTGACC  300
     ACGTCCAGCT ACACCCTGTC CTGCAAGCTG CACACTTTCC TCTTCGAGGC CTGCAGCTAC  360
     GCTACGCTGC TGCACGTGCT GACGCTCAGC TTTGAGCGCT ACATCGCCAT CTGTCACCCC  420
     TTCAGGTACA AGGCTGTGTC GGGACCTTGC CAGGTGAAGC TGCTGATTGG CTTCGTCTGG  480
     GTCACCTCCG CCCTGGTGGC ACTGCCCTTG CTGTTTGCCA TGGGTACTGA GTACCCCCTG  540
     GTGAACGTGC CCAGCCACCG GGGTCTCACT TGCAACCGCT CCAGCACCCG CCACCACGAG  600
     CAGCCCGAGA CCTCCAATAT GTCCATCTGT ACCAACCTCT CCAGCCGCTG GACCGTGTTC  660
     CAGTCCAGCA TCTTCGGCGC CTTCGTGGTC TACCTCGTGG TCCTGCTCTC CGTAGCCTTC  720
     ATGTGCTGGA ACATGATGCA GGTGCTCATG AAAAGCCAGA AGGGCTCGCT GGCCGGGGGC  780
     ACGGGGCCTC CGCAGCTGAG GAAGTCCGAG AGCGAAGAGA GCGGACCGC CAGGAGGCAG  840
     ACCATCATCT TCCTGAGGCT GATTGTTGTG ACATTGGCCG TATGCTGGAT GCCCAACCAG  900
     ATTCGGAGGA TCATGGCTGC GGCCAAACCC AAGCACGACT GGACGAGGTC CTACTTCCGG  960
     GCGTACATGA TCCTCCTCCC CTTCTCGGAG ACGTTTTTCT ACCTCAGCTC GGTCATCAAC 1020
     CCGCTCCTGT ACACGGTGTC CTCGCAGCAG TTTCGGCGGG TGTTCGTGCA GGTGCTGTGC 1080
     TGCCGCCTGT CGCTGCAGCA CGCCAACCAC GAGAAGCGCC TGCGCGTACA TGCGCACTCC 1140
     ACCACCGACA GCGCCCGCTT TGTGCAGCGC CCGTTGCTCT TCGCGTCCCG GCGCCAGTCC 1200
     TCTGCAAGGA GAACTGAGAA GATTTTCTTA AGCACTTTTC AGAGCGAGCC CGAGCCCCAG 1260
     TCTAAGTCCC AGTCATTGAG TCTCGAGTCA CTAGAGCCCA ACTCAGGCGC GAAACCAGCC 1320
     AATTCTGCTG CAGAGAATGG TTTTCAGGAG CATGAAGTTT GA
```

Seq ID NO: 34 Protein sequence
Protein Accession #: NP_001499.1

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     MASPSLPGSD CSQIIDHSHV PEFEVATWIK ITLILVYLII FVMGLLGNSV TIRVTQVLQK   60
     KGYLQKEVTD BMVSLACSDI LVFLIGMPME FYSIIWNPLT TSSYTLSCKL HTFLFEACSY  120
     ATLLHVLTLS FERYIAICHP FRYKAVSGPC QVKLLIGFVW VTSALVALPL LFAMGTEYFL  180
     VNVPSHRGLT CNRSSTRHHE QPETSNMSIC TNLSSRWTVF QSSIFGAFVV YLVVLLSVAF  240
     MCWNMMQVLM KSQKGSLAGG TRPPQLRKSE SEESRTARRQ TIIFLRLIVV TLAVCWMPNQ  300
     IRRIMAAAKP KEDWTRSYFR AYMILLLPFSE TFFYLSSVIN PLLYTVSSQQ FRRVFVQVLC  360
     CRLSLQHANE EERLRVHAHS TTDSARFVQR PLLFASRRQS SARRTEKIFL STFQSEAEPQ  420
     SKSQSLSLES LEPNSGAKPA NSAAENGPQE HEV
```

Seq ID NO: 35 DNA sequence
Nucleic Acid Accession #: NM_006475.1
Coding sequence: 28..2538

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     AACAGAACTG CAACGGAGAG ACTCAAGATG ATTCCCTTTT TACCCATGTT TTCTCTACTA   60
     TTGCTGCTTA TTGTTAACCC TATAAACGCC AACAATCATT ATGACAAGAT CTTGGCTCAT  120
     AGTCGTATCA GGGGTCGGGA CCAAGGCCCA AATGTCTGTG CCCTTCAACA GATTTTGGGC  180
     ACCAAAAAGA AATACTTCAG CACTTGTAAG AACTGGTATA AAAAGTCCAT CTGTGGACAG  240
     AAAACGACTG TTTTATATGA ATGTTGCCCT GGTTATATGA GAATGGAAGG AATGAAAGGC  300
     TGCCCAGCAG TTTTGCCCAT TGACCATGTT TATGGCACTC TGGGCATCGT GGGAGCCACC  360
     ACAACGCAGC GCTATTCTGA CGCCTCAAAA CTGAGGGAGG AGATCGAGGG AAAGGGATCC  420
     TTCACTTACT TTGCACCGAG TAATGAGGCT TGGGACAACT TGGATTCTGA TATCCGTAGA  480
     GGTTTGGAGA GCAACGTGAA TGTTGAATTA CTGAATGCTT TACATAGTCA CATGATTAAT  540
     AAGAGAATGT TGACCAAGGA CTTAAAAAAT GGCATGGATT ATTCCTTCAAT GTATAACAAT  600
     TTGGGGCTTT TCATTAACCA TTATCCTAAT GGGGTTGTCA CTGTTAATTG TGCTCGAATC  660
     ATCCATGGGA ACCAGATTGC AACAAATGGT GTTGTCCATG TCATTGACCG TGTGCTTACA  720
     CAAATGGTA CCTCAATTCA AGACTTCATT GAAGCAGAAG ATGACCTTTC ATCTTTAGA  780
     GCAGCTGCCA TCACATCGGA CATATTGGAG GCCCTTGGAA GAGACGGTCA CTTCACACTC  840
     TTTGCTCCCA CCAATGAGGC TTTTGAGAAA CTTCCACGAG GTGTCCTAGA AAGGTCATG  900
     GGAGACAAAG TGGCTTCCGA AGCTCTTATG AAGTACCACA TCTTAAATAC TCTCCAGTGT  960
     TCTGAGTCTA TTATGGGAGG AGCAGTCTTT GAGACGCTGG AAGGAAATAC AATTGAGATA 1020
```

```
GGATGTGACG GTGACAGTAT AACAGTAAAT GGAATCAAAA TGGTGAACAA AAAGGATATT  1080
GTGACAAATA ATGGTGTGAT CCATTTGATT GATCAGGTCC TAATTCCTGA TTCTGCCAAA  1140
CAAGTTATTG AGCTGGCTGG AAAACAGCAA ACCACCTTCA CGGATCTTGT GGCCCAATTA  1200
GGCTTGGCAT CTGCTCTGAG GCCAGATGGA GAATACACTT TGCTGGCACC TGTGAATAAT  1260
GCATTTTCTG ATGATACTCT CAGCATGGTT CAGCGCCTCC TTAAATTAAT TCTGCAGAAT  1320
CACATATTGA AAGTAAAAGT TGGCCTTAAT GAGCTTTACA ACGGGCAAAT ACTGGAAACC  1380
ATCGGAGGCA AACAGCTCAG AGTCTTCGTA TATCGTACAG CTGTCTGCAT TGAAAATTCA  1440
TGCATGGAGA AAGGGAGTAA GCAAGGGAGA AACGGTGCGA TTCACATATT CCGCGAGATC  1500
ATCAAGCCAG CAGAGAAATC CCTCCATGAA AAGTTAAAAC AAGATAAGCG CTTTAGCACC  1560
TTCCTCAGCC TACTTGAAGC TGCAGACTTG AAAGAGCTCC TGACACAACC TGGAGACTGG  1620
ACATTATTTG TGCCAACCAA TGATGCTTTT AAGGGAATGA CTAGTGAAGA AAAAGAAATT  1680
CTGATACGGG ACAAAAATGC TCTTCAAAAC ATCATTCTTT ATCACCTGAC ACCAGGAGTT  1740
TTCATTGGAA AAGGATTTGA ACCTGGTGTT ACTAACATTT TAAAGACCAC ACAAGGAAGC  1800
AAAATCTTTC TGAAAGAAGT AAATGATACA CTTCTGGGTGA ATGAATTGAA ATCAAAAGAA  1860
TCTGACATCA TGACAACAAA TGGTGTAATT CATGTTGTAG ATAAACTCCT CTATCCAGCA  1920
GACACACCTG TTGGAAATGA TCAACTGCTG GAAATACTTA ATAAATTAAT CAAATACATC  1980
CAAATTAAGT TTGTTCGTGG TAGCACCTTC AAAGAAATCC CCGTGACTGT CTATACAACT  2040
AAAATTATAA CCAAAGTTGT GGAACCAAAA ATTAAAGTGA TTGAAGGCAG TCTTCAGCCT  2100
ATTATCAAAA CTGAAGGACC CACACTAACA AAAGTCAAAA TTGAAGGTGA ACCTGAATTC  2160
AGACTGATTA AAGAAGGTGA AACAATAACT GAAGTGATCC ATGGAGAGCC AATTATTAAA  2220
AAATACACCA AAATCATTGA TGGAGTGCCT GTGGAAATAA CTGAAAAAGA GACACGAGAA  2280
GAACGAATCA TTACAGGTCC TGAAATAAAA TACACTAGGA TTTCTACTGG AGGTGGAGAA  2340
ACAGAAGAAA CTCTGAAGGAA ATTGTTACAA GAAGAGGTCA CCAAGGTCAC CAAATTCATT  2400
GAAGGTGGTG ATGGTCATTT ATTTGAAGAT GAAGAAATTA AAAGACTGCT TCAGGGAGAC  2460
ACACCGGTGA GGAAGTTGCA AGCCAACAAA AAAGTTCAAG GTTCTAGAAG ACGATTAAGG  2520
GAAGGTCGTT CTCAGTGAAA ATCCAAAAAC CAGAAAAAAA TGTTTATACA ACCCTAAGTC  2580
AATAACCTGA CCTTAGAAAA TTGTGAGAGC CAAGTTGACT TCAGGAACTG AAACATCAGC  2640
ACAAAGAAGC AATCATCAAA TAATTCTGAA CACAAATTTA ATATTTTTTT TTCTGAATGA  2700
GAAACATGAG GGAAATTGTG GAGTTAGCCT CCTGTGGTAA AGGAATTGAA GAAAATATAA  2760
CACCTTACAC CCTTTTTCAT CTTGACATTA AAAGTTCTGG CTAACTTTGG AATCCATTAG  2820
AGAAAAATCC TTGTCACCAG ATTCATTACA ATTCAAATCG AAGAGTTGTG AACTGTTATC  2880
CCATTGAAAA GACCGAGCCT TGTATGTATG TTATGGATAC ATAAAATGCA CGCAAGCCAT  2940
TATCTCTCCA TGGGAAGCTA AGTTATAAAA ATAGGTGCTT GGTGTACAAA ACTTTTTATA  3000
TCAAAAGGCT TTGCACATTT CTATATGAGT GGGTTTACTG GTAAATTATG TTATTTTTTA  3060
CAACTAATTT TGTACTCTCA GAATGTTTGT CATATGCTTC TTGCAATGCA TATTTTTTAA  3120
TCTCAAACGT TTCAATAAAA CCATTTTTCA GATATAAAGA GAATTACTTC AAATTGAGTA  3180
ATTCAGAAAA ACTCAAGATT TAAGTTAAAA AGTGGTTTGG ACTTGGGAA
```

Seq ID NO: 36 Protein sequence
Protein Accession #: NP_006466.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MIPFLPMFSL LLLLIVNPIN ANNHYDKILA HSRIRGRDQG PNVCALQQIL GTKRKYFSTC   60
KNWYKKSICG QKTTVLYECC PGYMRMEGMK GCPAVLPIDH VYGTLGIVGA TTTQRYSDAS  120
KLREEIEGKG SFTYFAPSNE AWDNLDSDIR RGLESNVNVE LLNALHSHMI NKRMLTKULK  180
NGMIIPSMYN NLGLFINHYP NGVVTVNCAR IIHGNQIATN GVVHVIDRVL TQIGTSIQDF  240
IEAEDDLSSF RAAAITSDIL EALGRDGHFT LFAPTNEAFE KLPRGVLERF MGDKVASEAL  300
MKYHILNTLQ CSESIMGGAV FETLEGNTIE IGCDGDSITV NGIKMVNKKD IVTNNGVIHL  360
IDQVLIPDSA KQVIELAGKQ QTTFTDLVAQ LGLASALRPD GEYTLLAPVN NAFSDDTLSM  420
VQRLLKLILQ NHILKVKVGL NELYNGQILE TIGGKQLRVF VYRTAVCIEN SCMEKGSKQG  480
RNGAIHIFRE IIKPAEKSLH EKLKQDKRFS TFLSLLEAD LKELLTQPGD WTLFVPTNDA  540
FKGMTSEEKE ILIRDKNALQ NIILYHLTPG VFIGKGFEPG VTNILKTIQG SKIFLKEVND  600
TLLVNELKSR ESDIMTTNGV IHVVDKLLYP ADTPVGNDQL LEILNKLIKY IQIKFVRGST  660
FKEIPVTVYT TKIITKVVEP KIKVIEGSLQ PIIKTEGPTL TKVKIEGEPE FRLIKEGETI  720
TEVIHGEPII KKYTKIIDGV PVEITEKETR EERIITGPEI KYTRISTGGG ETEETLKKLL  780
QEEVTKVTKF IEGGDGHLFE DEEIKRLLQG DTPVRKLQAN KKVQGSRRRL REGRSQ
```

Seq ID NO: 37 DNA sequence
Nucleic Acid Accession #: NM_002416
Coding sequence: 40..417

```
1          11         21         31         41         51
|          |          |          |          |          |
ATCCAATACA GGAGTGACTT GGAACTCCAT TCTATCACTA TGAAGAAAAG TGGTGTTCTT   60
TTCCTCTTGG GCATCATCTT GCTGGTTCTG ATTGGAGTGG AAGGAACCCC AGTAGTGAGA  120
AAGGGTCGCT GTTCCTGCAT CAGCACCAAC CAAGGGACTA TCCACCTACA ATCCTTGAAA  180
GACCTTAAAC AATTTGCCCC AAGCCCTTCC TGCGAGAAAA TTGAAATCAT TGCTACACTG  240
AAGAATGGAG TTCAAACATG TCTAAACCCA GATTCAGCAG ATGTGAAGGA ACTGATTAAA  300
AAGTGGGAGA AACAGGTCAG CCAAAAGAAA AAGCAAAAGA ATGGGAAAAA ACATCAAAGA  360
AAGAAAGTTC TGAAAGTTCG AAAATCTCAA CGTTCTCGTC AAAAGAAGAC TACATAAGAG  420
ACCACTTCAC CAATAAGTAT TCTGTGTTAA AAATGTTCTA TTTTAATTAT ACCGCTATCA  480
TTCCAAAGGA GGATGGCATA TAATACAAAG GCTTATTAAT TTGACTAGAA AATTTAAAAC  540
ATTACTCTGA AATTGTAACT AAAGTTAGAA AGTTGATTTT AAGAATCCAA ACGTTAAGAA  600
TTGTTAAAGG CTATGATTGT CTTTGTTCTT CTACCACCCA CCAGTTGAAT TTCATCATGC  660
TTAAGGCCAT GATTTTAGCA ATACCCATGT CTACACAGAT GTTCACCCAA CCACATCCCA  720
CTCACAACAG CTGCCTGGAA GAGCAGCCCT AGGCTTCCAC GTACTGCAGC CTCCAGAGAG  780
TATCTGAGGC ACATGTCAGC AAGTCCTAAG CCTGTTAGCA TGCTGGTGAG CCAAGCAGTT  840
TGAAATTGAG CTGGACCTCA CCAAGTCGCT GTGGCCATCA ACCTCTGTAT TTGAATCAGC  900
CTACAGGCCT CACACACAAT GTGTCTGAGA GATTCATGCT GATTGTTATT GGGTATCACC  960
ACTGGAGATC ACCAGTGTGT GGCTTTCAGA GCCTCCTTTC TGGCTTTGGA AGCCATGTGA 1020
TTCCATCTTG CCCGCTCAGG CTGACCACTT TATTTCTTTT TGTTCCCCTT TGCTTCATTC 1080
AAGTCAGCTC TTCTCCATCC TACCACAATG CAGTGCCTTT CTTCTCTCCA GTGCACCTGT 1140
CATATGCTCT GATTTATCTG AGTCAACTCC TTTCTCATCT TGTCCCCAAC ACCCCACAGA 1200
```

```
      AGTGCTTTCT TCTCCCAATT CATCCTCACT CAGTCCAGCT TAGTTCAAGT CCTGCCTCTT 1260
      AAATAAACCT TTTTGGACAC ACAAATTATC TTAAAACTCC TGTTTCACTT GGTTCAGTAC 1320
      CACATGGGTG AACACTCAAT GGTTAACTAA TTCTTGGGTG TTTATCCTAT CTCTCCAACC 1380
5     AGATTGTCAG CTCCTTGAGG GCAAGAGCCA CAGTATATTT CCCTGTTTCT TCCACAGTGC 1440
      CTAATAATAC TGTGGAACTA GGTTTAAATA ATTTTTTAAT TGATGTTGTT ATGGGAGGAA 1500
      TGGCAACCAG ACCATTGTCT CAGAGCAGGT GCTGGCTCTT TCCTGGCTAC TCCATGTTGG 1560
      CTAGCCTCTG GTAACCTCTT ACTTATTATC TTCAGGACAC TCACTACAGG GACCAGGGAT 1620
      GATGCAACAT CCTTGTCTTT TTATGACAGG ATGTTTGCTC AGCTTCTCCA ACAATAAGAA 1680
10    GCACGTGGTA AAACACTTGC GGATATTCTG GACTGTTTTT AAAAAAATAA CAGTTTACCG 1740
      AAAATCATAT AATCTTACAA TGAAAAGGAC TTTATAGATC AGCCAGTGAC CAACCTTTTC 1800
      CCAACCTATAC AAAAATTCCT TTTCCCGAAG GAAAAAGGGCT TTCTCAATAA GCCTCAGCTT 1860
      TCTAAGATCT AACAAGATAG CCACCGAGAT CCTTATCGAA ACTCATTTTA GGCAAATATG 1920
15    AGTTTTATTG TCCGTTTACT TGTTTCAGAG TTTGTATTGT GATTATCAAT TACCACACCA 1980
      TCTCCCATGA AGAAAGGGAA CGGTGAAGTA CTAAGCGCTA GAGGAAGCAG CCAAGTCGGT 2040
      TAGTGGAAGC ATGATTGGTG CCCAGTTAGC CTCTGCAGGA TGTGGAAACC TCCTTCCAGG 2100
      GGAGGTTCAG TGAATTGTGT AGGAGAGGTT GTCTGTGGCC AGAATTTAAA CCTATACTCA 2160
      CTTTCCCAAA TTGAATCACT GCTCACACTG CTGATGATTT AGAGTGCTGT CCGGTGGAGA 2220
20    TCCCACCCGA ACGTCTTATC TAATCATGAA ACTCCCTAGT TCCTTCATGT AACTTCCCTG 2280
      AAAAATCTAA GTGTTTCATA AATTTGAGAG TCTGTGACCC ACTTACCTTG CATCTCACAG 2340
      GTAGACAGTA TATAACTAAC AACCAAAGAC TACATATTGT CACTGACACA CACGTTATAA 2400
      TCATTTATCA TATATATACA TACATGCATA CACTCTCAAA GCAAATAATT TTTCACTTCA 2460
      AAACAGTATT GACTTGTATA CCTTGTAATT TGAAATATTT TCTTTGTTAA AATAGAATGG 2520
      TATCAATAAA TAGACCATTA ATCAG
```

25 Seq ID NO: 38 _Protein sequence_
    Protein Accession #: NP_002407

```
    1          11         21         31         41         51
    |          |          |          |          |          |
30  MKKSGVLFLL GIILLVLIGV QGTPVVRKGR CSCISTNQGT IHLQSLKDLK QFAPSPSCEK  60
    IEIIATLKNG VQTCLNPDSA DVKELIKKWE KQVSQKKKQK NGKKHQKKKV LKVRKSQRSR  120
    QKKTT
```

35 Seq ID NO: 39 _DNA sequence_
    Nucleic Acid Accession #: NM_006670
    Coding sequence: 85..1347

```
    1          11         21         31         41         51
    |          |          |          |          |          |
40  CCGGCTCGCG CCCTCCGGGC CCAGCCTCCC GAGCCTTCGG AGCGGGCGCC GTCCCAGCCC  60
    AGCTCCGGGG AAACGCGAGC CGCGATGCCT GGGGGGGTGCT CCCGGGGCCC CGCCGCCGGG  120
    GACGGGCGTC TGCGGCTGGC GCGACTAGCG CTGGTACTCC TGGGCTGGGT CTCCTCGTCT  180
    TCTCCCACCT CCTCGGCATC CTCCTTCTCC TCCTCGGCGC CGTTCCTGGC TTCCGCCGTG  240
    TCCGCCCAGC CCCCGGTCGCC GGACCAGTGC CCCGGCGCTGT GCGAGGCAGCG  300
    CGCACAGTCA AGTGCGTTAA CCGCAATCTG ACCGAGGTGC CCACGGACCT GCCCGCCTAC  360
    GTGCGCAACC TCTTCCTTAC CGGCAACCAG CTGGCCGTGC TCCCTGCCGG CGCCTTCGCC  420
    CGCCGCCGC CGCTGGCGGA GCTGGCCGCG CTCAACCTCA GCGGCAGCCG CCTGGACGAG  480
    GTGCGCGCGG GCGCCTTCGA GCATCTGCCC AGCCTGCGCC AGCTCGACCT CAGCCACAAC  540
    CCACTGGCCG ACCTCAGTCC CTTCGCTTTC TCGGGCAGCA ATGCCAGCGT CTCGGCCCCC  600
    AGTCCCCTTG TGGAACTGAT CCTGAACCAC ATCGTGCCCC CTGAAGATGA GCGGCAGAAC  660
    CGGAGCTTCG AGGGCATGGT GGTGGCGGCC CTGCTGGCGG GCCGTGCACT GCAGGGGCTC  720
    CGCCGCTTGG AGCTGGCCAG CAACCACTTC CTTTACCTGC CGCGGGATGT GCTGGCCCAA  780
    CTGCCCAGCC TCAGGCACCT GGACTTAAGT AATAATTCGC TGGTGAGCGT GACCTACGTG  840
    TCCTTCCGCA ACCTGACACA TCTAGAAAGC CTCCACCTGG AGGACAATGC CCTCAAGGTC  900
    CTTCACAATG GCACCCTGGC TGAGTTGCAA GGTCTACCCC ACATTAGGGT TTTCCTGGAC  960
    AACAATCCCT GGGTCTGCGA CTGCCACATG GCAGACATGG TGACCTGGCT CAAGGAAACA  1020
    GAGGTAGTGC AGGGCAAAGA CCGGCTCACC TGTGCATATC CGGAAAAAAT GAGGAATCGG  1080
    GTCCTCTTGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTCC CCCATCCCTG  1140
    CAAACCTCTT ATGTCTTCCT GGGTATTGTT TTAGCCCTGA TAGGCGCTAT TTTCCTCCTG  1200
    GTTTTGTATT TGAACCGCAA GGGGGATAAA AAGTGGATGC ATAACATCAG AGATGCCTGC  1260
    AGGGATCACA TGGAAGGGTA TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAACA  1320
    AACCTCAGTT CTAACTCGGA TGTCTGAGAA ATATTAGAGG ACAGACCAAG GACAACTCTG  1380
    CATGAGATGT AGACTTAAGC TTTATCCCTA CTAGGCTTGC TCCACTTTCA TCCTCCACTA  1440
    TAGATACAAC GGACTTTGAC TAAAAGCAGT GAAGGGGATT TGCTTCCTTG TTATGTAAAG  1500
    TTTCTCGGTG TGTTCTGTTA ATGTAAGACG ATGAACAGTT GTGTATAGTG TTTTACCCTC  1560
    TTCTTTTTCT TGGAACTCCT CAACACGTAT GGAGGGATTT TTCAGGTTTC AGCATGAACA  1620
    TGGGCTTCTT GCTGTCTGTC TCTCTCTCAG TACAGTTCAA GGTGTAGCAA GTGTACCCAC  1680
    ACAGATAGCA TTCAACAAAA GCTGCCTCAA CTTTTTCGAG AAAAATACTT TATTCATAAA  1740
    TATCAGTTTT ATTCTCATGT ACCTAAGTTG TGGAGAAAAT AATTGCATC TATAAACTGC  1800
    CTGCAGACGT TAGCAGGCTC TTCAAAATAA CTCCATGGTG CACAGGAGCA CCTGCATCCA  1860
    AGAGCATGCT TACATTTTAC TGTTCTGCAT ATTACAAAAA ATAACTTGCA ACTTCATAAC  1920
    TTCTTTGACA AAGTAAATTA CTTTTTTTGAT TGCAGTTTAT ATGAAAATGT ACTGATTTTT  1980
    TTTTAATAAA CTGCATCAG ATCCAACCGA CTGAATTGTT AAAAAAAAAA AAAAATAAAG  2040
    ATTCTTAAAA GAA
```

75 Seq ID NO: 40 _Protein sequence_
    Protein Accession #: NP_006661

```
    1          11         21         31         41         51
    |          |          |          |          |          |
80  MPGGCSRGPA AGDGRLRLAR LALVLLGWVS SSSPTSSASS PSSSAPFLAS AVSAQPPLPD  60
    QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRNLFLTG NQLAVLPAGA FARRPPLAEL  120
    AALNLSGSRL DEVRAGAFEH LPSLRQLDLS HNPLADLSPF APSGSNASVS APSFLVELIL  180
    NHIVPPEDER QNRSPEGMVV AALLAGRALQ GLRRLELASN HFLYLPRDVL AQLPSLRHLD  240
```

```
LSNNSLVSLT YVSFRNLTHL ESLHLEDNAL KVLHNGTLAE LQGLPHIRVF LDNNPWVCDC   300
EMADMVTWLK ETEVVQGKDR LTCAYPEKMR NRVLLELNSA DLDCDPILPP SLQTSYVFLG   360
IVLALIGAIF LLVLYLNRKG IKKWMHNIRD ACRDHMEGYH YRYEINADPR LTNLSSNSDV
```

Seq ID NO: 41 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..927

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCCTGGGG GGTGCTCCCG GGGCCCCGCC GCCGGGGACG GGCGTCTGCG GCTGGCGCGA   60
CTAGCGCTGG TACTCCTGGG CTGGGTCTCC TCGTCTTCTC CCACCTCCTC GGCATCCTCC   120
TTCTCCTCCT CGGCGCCGTT CCTGGCTTCC GCCGTGTCCG CCCAGCCCGC GCTGCCGGAC   180
CAGTGCCCCG CGCTGTGCGA GTGCTCCGAG GCAGCGCGCA CAGTCAAGTG CGTTAACCGC   240
AATCTGACCG AGGTGCCCAC GGACCTGCCC GCCTACGTGC GCAACCTCTT CCTTACCGGC   300
AACCAGCTGG CCAGCAACCA CTTCCTTTAC CTGCCGCGGG ATGTGCTGGC CCAACTGCCC   360
AGCCTCAGGC ACCTGGACTT AAGTAATAAT TCGCTGGTGA GCCTGACCTA CGTGTCCTTC   420
CGCAACCTGA CACATCTAGA AAGCCTCCAC CTGGAGGACA ATGCCCTCAA GGTCCTTCAC   480
AATGGCACCC TGGCTGAGTT GCAAGGTCTA CCCCACATTA GGGTTTTCCT GGACAACAAT   540
CCCTGGGTCT GCGACTGCCA CATGGCAGAC ATGGTGACCT GGCTCAAGGA AACAGAGGTA   600
GTGCAGGGCA AAGACCGGCT CACCTGTGCA TATCCGGAAA AAATGAGGAA TCGGGTCCTC   660
TTGGAACTCA ACAGTGCTGA CCTGGACTGT GACCCGATTC TTCCCCCATC CCTGCAAACC   720
TCTTATGTCT TCCTGGGTAT TGTTTTAGCC CTGATAGGCG CTATTTTCCT CCTGGTTTTG   780
TATTTGAACC GCAAGGGGAT AAAAAAGTGG ATGCATAACA TCAGAGATGC CTGCAGGGAT   840
CACATGGAAG GGTATCATTA CAGATATGAA ATCAATGCGG ACCCCAGATT AACAAACCTC   900
AGTTCTAACT CGGATGTCCT CGAGTGA
```

Seq ID NO: 42 Protein sequence
Protein Accession #: Eos sequence

```
1          11         21         31         41         51
|          |          |          |          |          |
MPGGCSRGPA AGDGRLRLAR LALVLLGWVS SSSPTSSASS FSSSAPFLAS AVSAQPPLPD   60
QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRNLFLTG NQLASNHFLY LPRDVLAQLP   120
SLRHLDLSNN SLVSLTYVSF RNLTHLESLH LEDNALKVLH NGTLAELQGL PHIRVFLDNN   180
PWVCDCHMAD MVTWLKETEV VQGKDRLTCA YPEKMRNRVL LELNSADLDC DPILPPSLQT   240
SYVFLGIVLA LIGAIFLLVL YLNRKGIKKW MHNIRDACRD HMEGYHYRYE INADPRLTNL   300
SSNSDVLE
```

Seq ID NO: 43 DNA sequence
Nucleic Acid Accession #: NM_058173
Coding sequence: 68..340

```
1          11         21         31         41         51
|          |          |          |          |          |
AGCGCCTTGC CTTCTCTTAG GCTTTGAAGC ATTTTTGTCT GTGCTCCCTG ATCTTCAGGT   60
CACCACCATG AAGTTCTTAG CAGTCCTGGT ACTCTTGGGA GTTTCCATCT TTCTGGTCTC   120
TGCCCAGAAT CCGACAACAG CTGCTCCAGC TGACACGTAT CCAGCTACTG GTCCTGCTGA   180
TGATGAAGCC CCTGATGCTG AAACCACTGC TGCTGCAACC ACTGCGACCA CTGCTGCTCC   240
TACCACTGCA ACCACCGCTG CTTCTACCAC TGCTCGTAAA GACATTCCAG TTTTACCCAA   300
ATGGGTTGGG GATCTCCCGA ATGGTAGAGT GTGTCCCTGA GATGGAATCA GCTTGAGTCT   360
TCTGCAATTG GGTCACAACT ATTCATGCTT CCTGTGATTT CATCCAACTA CTTACCTTGC   420
CTACGATATC CCCTTTATCT CTAATCAGTT TATTTTCTTT CAAATAAAAA ATAACTATGA   480
GCGAGCTAAC AT
```

Seq ID NO: 44 Protein sequence
Protein Accession #: NP_477521

```
1          11         21         31         41         51
|          |          |          |          |          |
MKFLAVLVLL GVSIFLVSAQ NPTTAAPADT YPATGPADDE APDAETTAAA TTATTAAPTT   60
ATTAASTTAR KDIPVLPKWV GDLPNGRVCP
```

Seq ID NO: 45 DNA sequence
Nucleic Acid Accession #: NM_000095.1
Coding sequence: 26..2299

```
1          11         21         31         41         51
|          |          |          |          |          |
CAGCACCCAG CTCCCCGCCA CCGCCATGGT CCCCGACACC GCCTGCGTTC TTCTGCTCAC   60
CCTGGCTGCC CTCGGCGCGT CCGGACAGGG CCAGAGCCCG TTGGGCTCAG ACCTGGGCCC   120
GCAGATGCTT CGGGAACTGC AGGAAACCAA CGCGGCGCTG CAGGACGTGC GGGACTGGCT   180
GCGGCAGCAG GTCAGGGAGA TCACGTTCCT GAAAAACACG GTGATGGAGT GTGACGCGTG   240
CGGGATGCAG CAGTCAGTAC GCACCGGCCT ACCCAGCGTG CGGCCCCTGC TCCACTGCGC   300
GCCCGGCTTC TGCTTCCCCG GCGTGGCCTG CATCCAGACG GAGAGCGGCG GCCGCTGCGG   360
CCCCTGCCCC GCGGGCTTCA CGGGCAACGG CTCGCACTGC ACCGACGTCA ACGAGTGCAA   420
CGCCCACCCC TGCTTCCCCC GAGTCCGCTG TATCAACACC AGCCCGGGGT TCCGCTGCGA   480
GGCTTGCCCG CCGGGGTACA GCGGCCCCAC CCACCAGGGC GTGGGGCTGG CTTTCGCCAA   540
GGCCAACAAG CAGGTTTGCA CGGACATCAA CGAGTGTGAG ACCGGGCAAC ATAACTGCGT   600
CCCCAACTCC GTGTGCATCA ACACCCGGGG CTCCTTCCAG TGCGGCCCGT GCCAGCCCGG   660
CTTCGTGGGC GACCAGGCGT CGGCTGCCA GCGCGGCGCA CAGCGCTTCT GCCCGGACGG   720
CTCGCCCAGC GAGTGCCACG AGCATGCAGA CTGCGTCCTA GAGCGCGGATG GCTCGCGGTC   780
GTGCGTGTGT CGCGTTGGCT GGGCCGGCAA CGGGATCCTC TGTGGTCGCG ACACTGACCT   840
AGACGGCTTC CGGGACGAGA AGCTGCGCTG CCCGGAGCCG CAGTGCCGTA AGGACAACTG   900
```

```
        CGTGACTGTG CCCAACTCAG GGCAGGAGGA TGTGGACCGC GATGGCATCG GAGACGCCTG   960
        CGATCCGGAT GCCGACGGGG ACGGGGTCCC CAATGAAAAG GACAACTGCC CGCTGGTGCG  1020
        GAACCCAGAC CAGCGCAACA CGGACGAGGA CAAGTGGGGC GATGCGTGCG ACAACTGCCG  1080
        GTCCCAGAAG AACGACGACC AAAAGGACAC AGACCAGGAC GGCCGGGGCG ATGCGTGCGA  1140
  5     CGACGACATC GACGGCGACC GGATCCGCAA CCAGGCCGAA AACTGCCCTA GGGTACCCAA  1200
        CTCAGACCAG AAGGACAGTG ATGGCGATGG TATAGGCGGAT GCCTGTGACA ACTGTCCCCA  1260
        GAAGAGCAAC CCGGATCAGG CGGATGTGGA CCACGACTTT GTGGGAGATG CTTGTGACAG  1320
        CGATCAAGAC CAGGATGGAG ACGGACATCA GGACTCTCGG GACAACTGTC CCACGGTGCC  1380
        TAACAGTGCC CAGGAGGACT CAGACCACGA TGGCCAGGGT GATGCCTGCG ACGACGACGA  1440
  10    CGACAATGAC GGAGTCCCTG ACAGTCGGGA CAACTGCCGC CTGGTGCCTA ACCCCGGCCA  1500
        GGAGGACGCG GACAGGGACG GCGTGGGCGA CGTGTGCCAG GACGACTTTG ATGCAGACAA  1560
        GGTGGTAGAC AAGATCGACG TGTGTCCGGA GAACGCTGAA GTCACGCTCA CCGACTTCAG  1620
        GGCCTTCCAG ACAGTCGTGC TGGACCCGGA GGGTGACGCG CAGATTGACC CCAACTGGGT  1680
        GGTGCTCAAC CAGGGAAGGG AGATCGTGCA GACAATGAAC AGCGCACCCAG GCCTGGCTGT  1740
  15    GGGTTACACT GCCTTCAATG GCGTGGACTT CGAGGGCACG TTCCATGTGA ACACGGTCAC  1800
        GGATGACGAC TATGCGGGCT TCATCTTTGG CTACCAGGAC AGCTCCAGCT TCTACGTGGT  1860
        CATGTGGAAG CAGATGGAGC AAACGTATTG GCAGGCGAAC CCCTTCCGTG CTGGTGGCCGA  1920
        GCCTGGCATC CAACTCAAGG CTGTGAAGTC TTCCACAGGC CCCGGGGAAC AGCTGCGGAA  1980
        CGCTCTGTGG CATACAGGAG ACACAGAGTC CCAGGTGCGG CTGCTGTGGA AGGACCCGCG  2040
  20    AAACGTGGGT TGGAAGGACA AGAAGTCCTA TCGTTGGTTC CTGCAGCACC GGCCCCAAGT  2100
        GGGCTACATC AGGGTGCGAT TCTATGAGGG CCCTGAGCTG GTGGCCGACA GCAACGTGGT  2160
        CTTGGACACA ACCATGCGGG GTGGCCGCCT GGGGGTCTTC TGCTTCTCCC AGGAGAACAT  2220
        CATCTGGGCC AACCTGCGTT ACCGCTGCAA TGACACCATC CCAGAGGACT ATGAGACCCA  2280
        TCAGCTGCGG CAAGCCTAGG GACCAGGGTG AGGACCCGCC GGATGACAGC CACCCTCACC  2340
  25    GCGGCTGGAT GGGGGCTCTG CACCCCAGCCC AAGGGGTGGC CGTCCTGAGG GGGAAGTGAG  2400
        AAGGGCTCAG AGAGGACAAA ATAAAGTGTG TGTGCAGGG
```

```
Seq ID NO: 46 Protein sequence
Protein Accession #: NP_000086.1
```

```
  30
        1          11         21         31         41         51
        |          |          |          |          |          |
        MVPDTACVLL LTLAALGASG QGQSPLGSDL GPQMLRELQE TNAALQDVRD WLRQQVREIT   60
        FLRNTVMECD ACGMQQSVRT GLPSVRPLLH CAPGFCFPGV ACIQTESGGR CGPCPAGFTG  120
  35    NGSHCTDVNE CNAHPCFPRV RCINTSFGFR CEACPPGYSG PTHQGVGLAF AKANKQVCTD  180
        INECETGQHN CVPNSVCINT RGSFQCGPCQ PGFVGDQASG CQRGAQRFCP DGSPSECHEH  240
        ADCVLERDGS RSCVCRVGWA GNGILCGRDT DLDGFPDEKL RCPEPQCRKD NCVTVPNSGQ  300
  40    EDVDRDGIGD ACDPDADGDG VPNEKDNCPL VRNPDQRNTD EDKWGDACDN CRSQKNDDQK  360
        DTDQDGRGDA CDDDIDGDRI RNQADNCPRV PNSDQKDSDG DGIGDACDNC PQKSNPDQAD  420
        VDHDPVGDAC DSDQDQDGDG HQDSRDNCPT VPNSAQEDSD HDGQGDACDD DDDNDGVPDS  480
        RDNCRLVPNP GQEDADRDGV GDVCQDDFDA DKVVDKIDVC PENAEVTLTD FRAFQTVVLD  540
  45    PEGDAQIDPN WVVLNQGREI VQTMNSDPGL AVGYTAFNGV DFEGTFHVNT VTDDDYAGFI  600
        FGYQDSSSFY VVMWKQMEQT YWQANPFRAV AEPGIQLKAV KSSTGPGEQL RNALWHTGDT  660
        ESQVRLLWKD PRNVGWKDKK SYRNFLQHRP QVGYIRVRFY EGPELVADSN VVLDTTMRGG  720
        RLGVPCFSQE NIIWANLRYR CNDTIPEDYE TEQLRQA
```

```
Seq ID NO: 47 DNA sequence
Nucleic Acid Accession #: NM_001565.1
Coding sequence: 67..363
```

```
  50
        1          11         21         31         41         51
        |          |          |          |          |          |
        GAGACATTCC TCAATTGCTT AGACATATTC TGAGCCTACA GCAGAGGAAC CTCCAGTCTC   60
        AGCACCATGA ATCAAACTGC GATTCTGATT TGCTGCCTTA TCTTTCTGAC TCTAAGTGGC  120
  55    ATTCAAGGAG TACCTCTCTC TAGACCCGTA CGCTGCACCT GCATCAGCAT TAGTAATCAA  180
        CCTGTTAATC CAAGGTCTTT AGAAAAACTT GAAATTATTC CTGCAAGCCA ATTTTGTCCA  240
        CGTGTTGAGA TCATTGCTAC AATGAAAAAG AAGGGTGAGA AGAGATGTCT GAATCCAGAA  300
        TCGAAGGCCA TCAAGAATTT ACTGAAAGCA GTTAGCAAGG AAATGTCTAA AAGATCTCCT  360
        TAAAACCAGA GGGGAGCAAA ATCGATGCAG TGCTTCCAAG GATGGACCAC ACAGAGGCTG  420
  60    CCTCTCCCAT CACTTCCCTA CATGGAGTAT ATGTCAAGCC ATAATTGTTC TTAGTTTGCA  480
        GTTACACTAA AAGGTGACCA ATGATGGTCA CCAAATCAGC TGCTACTACT CCTGTAGGAA  540
        GGTTAATGTT CATCATCCTA AGCTATTCAG TAATAACTCT ACCCTGGCAC TATAATGTAA  600
        GCTCTACTGA GGTGCTATGT TCTTAGTGGA TGTTCTGACC CTGCTTCAAA TATTTCCCTC  660
        ACCTTTCCCA TCTTCCAAGG GTACTAAGGA ATCGTTCTGC TTTGGGGTTT ATCAGAATTC  720
  65    TCAGAATCTC AAATAACTAA AAGGTATGCA ATCAAATCTG CTTTTTAAAG AATGCTCTTT  780
        ACTTCATGGA CTTCCACTGC CATCCTCCCA AGGGGCCCAA ATTCTTTCAG TGGCTACCTA  840
        CATACAATTC CAAACACATA CAGGAAGGTA GAAATATCTG AAAATGTATG TGTAAGTATT  900
        CTTATTTAAT GAAAGACTGT ACAAAGTATA AGTCTTAGAT GTATATATTT CCTATATTGT  960
        TTTCAGTGTA CATGGAATAA CATGTAATTA AGTACTATGT ATCAATGAGT AACAGGAAAA  1020
  70    TTTTAAAAAT ACAGATAGAT ATATGCTCTG CATGTTACAT AAGATAAATG TGCTGAATGG  1080
        TTTTCAAATA AAAATGAGGT ACTCTCCTGG AAATATTAAG
```

```
Seq ID NO: 48 Protein sequence
Protein Accession #: NP_001556.1
```

```
  75
        1          11         21         31         41         51
        |          |          |          |          |          |
        MNQTAILICC LIFLTLSGIQ GVPLSRTVRC TCISISNQPV NPRSLEKLEI IPASQFCPRV   60
        EIIATMKKKG EKRCLNPESK AIKNLLKAVS KEMSKRSP
  80
```

```
Seq ID NO: 49 DNA sequence
Nucleic Acid Accession #: XM_057014
Coding sequence: 143..874
```

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       GGGAGGGAGA GAGGCGCGCG GGTGAAAGGC GCATTGATGC AGCCTGCGGC GGCCTCGGAG    60
       CGCGGCGGAG CCAGACGCTG ACCACGTTCC TCTCCTCGGT CTCCTCCGCC TCCAGCTCCG   120
       CGCTGCCCGG CAGCCGGGAG CCATGCGACC CAGGGCCCCG GCCGCCTCCC CGCAGCGGCT   180
       CCGCGGCCTC CTGCTGCTCC TGCTGCTGCA GCTGCCCGCG CCGTCGAGCG CCTCTCGAGAT   240
       CCCCAAGGGG AAGCAAAAGG CGCAGCTCCG GCAGAGGGAG GTGGTGGACC TGTATAATGG   300
       AATGTGCTTA CAAGGGCCAG CAGGAGTGCC TGGTCGAGAC GGGAGCCCTG GGGCCAATGG   360
       CATTCCGGGT ACACCTGGGA TCCCAGGTCG GGATGGATTC AAAGGAGAAA AGGGGGAATG   420
       TCTGAGGGAA AGCTTTGAGG AGTCCTGGAC ACCCAACTAC AAGCAGTGTT CATGGAGTTC   480
       ATTGAATTAT GGCATAGATC TTGGGAAAAT TGCGGAGTGT ACATTTACAA AGATGCGTTC   540
       AAATAGTGCT CTAAGAGTTT TGTTCAGTGG CTCACTTCGG CTAAAATGCA GAAATGCATG   600
       CTGTCAGCGT TGGTATTTCA CATTCAATGG AGCTGAATGT TCAGGACCTC TTCCCATTGA   660
       AGCTATAATT TATTTGGACC AAGGAAGCCC TGAAATGAAT TCAACAATTA ATATTCATCG   720
       CACTTCTTCT GTGGAAGGAC TTTGTGAAGG AATTGGTGCT GGATTAGTGG ATGTTGCTAT   780
       CTGGGTTGGC ACTTGTTCAG ATTACCCAAA AGGAGATGCT TCTACTGGAT GGAATTCAGT   840
       TTCTCGCATC ATTATTGAAG AACTACCAAA ATAAATGCTT TAATTTTCAT TTGCTACCTC   900
       TTTTTTTATT ATGCCTTGGA ATGGTTCACT TAAATGACAT TTTAAATAAG TTTATGTATA   960
       CATCTGAATG AAAAGCAAAG CTAAATATGT TTACAGACCA AAGTGTGATT TCACACTGTT  1020
       TTTAAATCTA GCATTATTCA TTTTGCTTCA ATCAAAAGTG GTTTCAATAT TTTTTTTAGT  1080
       TGGTTAGAAT ACTTTCTTCA TAGTCACATT CTCTCAACCT ATAATTTGGA ATATTGTTGT  1140
       GGTCTTTTGT TTTTTCTCTT AGTATAGCAT TTTTAAAAAA ATATAAAAGC TACCAATCTT  1200
       TGTACAATTT GTAAATGTTA AGAATTTTTT TTATATCTGT TAAATAAAAA TTATTTCCAA  1260
       CAACCTTAAA AAAAAAAAAA AAAA
```

Seq ID NO: 50 Protein sequence
Protein Accession #: XP_057014

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       MRPQGPAASP QRLRGLLLLL LLQLPAPSSA SEIPKGKQKA QLRQREVVDL YNGMCLQGPA    60
       GVPGRDGSPG ANGIPGTPGI PGRDGFKGEK GECLRESFEE SWTPNYKQCS WSSLNYGIDL   120
       GKIAECTFTK MRSNSALRVL FSGSLRLKCR NACCQRWYFT FNGAECSGPL PIEAIIYLDQ   180
       GSPEMNSTIN IHRTSSVEGL CEGIGAGLVD VAIWVGTCSD YPKGDASTGW NSVSRIIIEE   240
       LPK
```

Seq ID NO: 51 DNA sequence
Nucleic Acid Accession #: NM_020974
Coding sequence: 81..3080

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       GGCGTCCGCG CACACCTCCC CGCGCCGCCG CCGCCACCGC CCGCACTCCG CCGCCTCTGC    60
       CCGCAACCGC TGAGCCATCC ATGGGGGTCG CGGGCCGCAA CCGTCCCGGG GCGGCCTGGG   120
       CGGTGCTGCT GCTGCTGCTG CTGCTGCCGC CACTGCTGCT GCTGGCGGGG GCCGTCCCGC   180
       CGGGTCGGGG CCGTGCCGCG GGGCCGCAGG AGGATGGTAGA TGAGTGTGCC CAAGGGCTAG   240
       ATGACTGCCA TGCCGACGCC CTGTGTCAGA ACACACCCAC CTCCTACAAG TGCTCCTGCA   300
       AGCCTGGCTA CCAAGGGGAA GGCAGGCAGT GTGAGGACAT CGATGAATGT GGAAATGAGC   360
       TCAATGGAGG CTGTGTCCAT GACTGTTTGA ATATTCCAGG CAATTATCGT TGCACTTGTT   420
       TTGATGGCTT CATGTTGGCT CATGACGGTC ATAATTGTCT TGATGTGGAC GAGTGCCTGG   480
       AGAACAATGG CGGCTGCCAG CATACCTGTG TCAACGTCAT GGGGAGCTAT GAGTGCTGCT   540
       GCAAGGAGGG GTTTTTCCTG AGTGACAATC AGCACACCTG CATTCACCGC TCGGAAGAGG   600
       GCCTGAGCTG CATGAATAAG GATCACGGCT GTAGTCACAT CTGCAAGGAG GCCCCAAGGG   660
       GCAGCGTCGC CTGTGAGTGC AGGCCTGGTT TTGAGCTGGC CAAGAACCAG AGAGACTGCA   720
       TCTTGACCTG TAACCATGGG AACGGTGGGT GCCAGCACTC CTGTGACGAT ACAGCCGATG   780
       GCCCAGAGTG CAGCTGCCAT CCACAGTACA AGATGCCAAG AGATGGGAGG AGCTGCCTTG   840
       AGCGAGAGGA CACTGTCCTG GAGGTGACAC AGAGCAACAC CACACATCAGTG GTGGATGGGG   900
       ATAAACGGGT GAAACGGCGG CTGCTCATGG AAACGTGTGC TGTCAACAAT GGAGGCTGTG   960
       ACCGCACCTG TAAGGATACT TGACACGGTG TCCACTGCAG TTGTCCTGTT GGATTCACTC  1020
       TCCAGTTGGA TGGGAAGACA TGTAAAGATA TTGATGAGTG CCAGACCCGC AATGGAGGTT  1080
       GTGATCATTT CTGCAAAAAC ATCGTGGGCA GTTTTGACTG CGGCTGCAAG AAAGGATTTA  1140
       AATTATTAAC AGATGAGAAG TCTTGCCAAG ATGTGGATGA GTGCTCTTTG GATAGGACCT  1200
       GTGACCACAG CTGCATCAAC ACCCCTGGCA CATTTGCTTG TGCTTGCAAC CGAGGGTACA  1260
       CCCTGTATGG CTTCACCCAC TGTGGAGACA CCAATGAGTG CAGCATCAAC AACGGAGGCT  1320
       GTCAGCAGGT CTGTGTGAAC ACAGTGGGCA GCTATGAATG CCAGTGCCAC CCTGGGTACA  1380
       AGCTCCACTG GAATAAAAAA GACTGTGTGG AAGTGAAGGG GCTCCTGCCC ACAAGTGTGT  1440
       CACCCCGTGT GTCCCTGCAC TGCGGTAAGA GTGGTGGAGG AGACGGGTGC TTCCTCAGAT  1500
       GTCACTCTGG CATTCACCTC TCTTCAGATG TCACCACCAT CAGGACAAGT GTAACCTTTA  1560
       AGCTAAATGA AGGCAAGTGT AGTTTGAAAA ATGCTGAACT GTTTCCCGAG GGTCTGCCAC  1620
       CAGCACTACC AGAGAAGCAC AGCTCAGTAA AAGAGAGCTT CCGCTACGTA AACCTTACAT  1680
       GCAGCTCTGG CAAGCAAGTC CCAGGAGCCC CTGGCCGACC AAGCACCCCT AAGGAAATGT  1740
       TTATCACTGT TGAGTTTGAG CTTGAAACTA ACCAAAAGCA GGTGACAGCT TCTTGTGACC  1800
       TGAGCTGCAT CGTAAAGCGA ACCGAGAAGC GGCTCCGTAA AGCCATCCGC ACGCTCAGAA  1860
       AGGCCGTCCA CAGGGAGCAG TTTCACCTCC AGCTCTCAGG CATGAACCTC GACGTGGCTA  1920
       AAAAGCCTCC CAGAAACATCT GAACGCCAGG CAGAGTCCTG TGGAGTGGGC CAGGGTCATG  1980
       CAGAAAACCA ATGTGTCAGT TGCAGGGCTG GGACCTATTA TGATGGAGCA CGAGAACGCT  2040
       GCATTTTATG TCCAAATGGA ACCTTCCAAA ATGAGGAAGG ACAAATGACT TGTGAACCAT  2100
       GCCCAAGACC AGGAAATTCT GGGGCCCTGA AGACCCCGAA AGCTTGGAAT ATGTCTGAAT  2160
       GTGGAGGTCT GTGTCAACCT GGTGAATATT CTGCAGATGG CTTTGCACCT TGCCAGCTCT  2220
       GTGCCCTGGG CACGTTCCAG CCTGAAGCTG GTCGAACTTC CTGCTTCCCC TGTGGAGGAG  2280
       GCCTTGCCAC CAAACATCAG GGAGCTACTT CCTTTCAGGA CTGTGAAACC AGAGTTCAAT  2340
       GTTCACCTGG ACATTTCTAC AACACCACCA CTCACCGATG TATTCGTTGC CCAGTGGGAA  2400
       CATACCAGCC TGAATTTGGA AAAAATAATT GTGTTTCTTG CCCAGGGAAT ACTACGACTG  2460
       ACTTTGATGG CTCCACAAAC ATAACCCAGT GTAAAAACAG AAGATGTGGA GGGGAGCTGG  2520
```

```
GAGATTTCAC TGGGTACATT GAATCCCCAA ACTACCCAGG CAATTACCCA GCCAACACCG 2580
AGTGTACGTG GACCATCAAC CCACCCCCCA AGCGCCGCAT CCTGATCGTG GTCCCTGAGA 2640
TCTTCCTGCC CATAGAGGAC GACTGTGGGG ACTATCTGGT GATGCGGAAA ACCTCTTCAT 2700
CCAATTCTGT GACAACATAT GAAACCTGCC AGACCTACGA ACGGCCCCATC GCCTTCACCT 2760
CCAGGTCAAA GAAGCTGTGG ATTCAGTTCA AGTCCAATGA AGGGAACCAC GCTAGAGGGT 2820
TCCAGGTCCC ATACGTGACA TATGATGAGG ACTACCAGGA ACTCATTGAA GACATAGTTC 2880
GAGATGGCAG GCTCTATGCA TCTGAGAACC ATCAGGAAAT ACTTAAGGAT AAGAAACTTA 2940
TCAAGGCTCT GTTTGATGTC CTGGCCCATC CCCAGAACTA TTTCAAGTAC ACAGCCCAGG 3000
AGTCCGAGA GATGTTCCA AGATCGTTCA TCCGATTGCT ACGTTCCAAA GTGTCCAGGT 3060
TTTTGAGACC TTACAAATGA CTCAGCCCAC GTGCCACTCA ATACAAATGT TCTGCTATAG 3120
GGTTGGTGGG ACAGAGCTGT CTTCCTTCTG CATGTCAGCA CAGTCGGGTA TTGCTGCCTC 3180
CCGTATCAGT GACTCATTAG AGTTCAATTT TTATAGATAA TACAGATATT TTGGTAAATT 3240
GAACTTGGTT TTTCTTTCCC AGCATCGTGG ATGTAGACTG AGAATGGCTT TGAGTGGCAT 3300
CAGCTTCTCA CTGCTGTGGG CGGATGTCTT GGATAGATCA CGGGCTGGCT GAGCTGGACT 3360
TTGGTCAGCC TAGGTGAGAC TCACCTGTCC TTCTGGGGTGT TTACTCCTCC TCAAGGAGTC 3420
TGTAGTGGAA AGGAGGCCAC AGAATAAGCT GCTTATTCTG AAACTTCAGC TTCCTCTAGC 3480
CCGGCCCTCT CTAAGGGAGC CCTCTGCACT CGTGTGCAGG CTCTGACCAG GCAGAACAGG 3540
CAAGAGGGGA GGGAAGGAGA CCCCTGCAGG CTCCCTCCAC CCACCTTGAG ACCTGGGAGG 3600
ACTCAGTTTC TCCACAGCCT TCTCCAGCCT GTGTGATACA AGTTTGATCC CAGGAACTTG 3660
AGTTCTAAGC AGTGCTCGTG AAAAAAAAAA GCAGAAAGAA TTAGAAATAA ATAAAAACTA 3720
AGCACTTCTG GAGACAT
```

Seq ID NO: 52 Protein sequence
Protein Accession #: NP_066025

```
1          11         21         31         41         51
|          |          |          |          |          |
MGVAGRNRPG AAWAVLLLLL LLPPLLLLAG AVPPGRGRAA GPQEDVDECA QGLDDCHADA   60
LCQNTPTSYK CSCKPGYQGE GRQCEDIDEC GNELNGGCVE DCLNIPGNYR CTCFDGFMLA  120
HDGHNCLDVD ECLENNGGCQ HTCVNVMGSY ECCCKEGFFL SDNQHTCIHR SEEGLSCMNK  180
DHGCSHICKE APRGSVACEC RPGFELAKNQ RDCILTCNHG NGGCQHSCDD TADGPECSCH  240
PQYKMHTDGR SCLEREDTVL EVTESNTTSV VDGDKRVKRR LLMETCAVNN GGCDRTCKDT  300
STGVHCSCPV GFTLQLDGKT CKDIDECQTR NGGCDHFCKN IVGSFDCGCK KGFKLLTDEK  360
SCQDVDECSL DRTCDHSCIN HPGTFACACN RGYTLYGFTH CGDTNECSIN NGGCQQVCVN  420
TVGSYECQCH PGYKLHWNKK DCVEVKGLLP TSVSPRVSLH CGKSGGGDGC FLRCHSGIHL  480
SSDVTTIRTS VTFKLNEGKC SLKNAELPPE GLRPALPEKH SSVKESFRYV NLTCSSGKQV  540
PGAPGRPSTP KEMFITVEFE LETNQKEVTA SCDLSCIVKR TKRLRKAIR TLRKAVHREQ  600
FHLQLSGMNL DVAKKPPRTS ERQAESCGVG QGHAENQCVS CRAGTYYDGA RERCILCPNG  660
TFQNEEGQMT CKPCPRPGNS GALKTPEAWN MSECGGLCQP GEYSADGFAP CQLCALGTFQ  720
PEAGRTSCFP CGGGLATKHQ GATSFQDCET RVQCSPGHFY NTTTHRCIRC PVGTYQPEFG  780
KNNCVSCPGN TTTDFDGSTN ITQCKNRRCG GELGDFTGYI ESPNYPGNYP ANTECTWTIN  840
PPPKRRILIV VPEIFLPIED DCGDYLVMRK TSSSNSVTTY ETCQTYERPI AFTSRSKKLW  900
IQFKSNEGNS ARGFQVPYVT YDEDYQHLIE DIVRDGRLYA SENHQEILKD KKLIKALFDV  960
LAHPQNYFKY TAQESREMFP RSFIRLLRSK VSRFLRPYK
```

Seq ID NO: 53 DNA sequence
Nucleic Acid Accession #: NM_014211
Coding sequence: 157..1479

```
1          11         21         31         41         51
|          |          |          |          |          |
GGGACAGGGC TGAGGATGAG GAGAACCCTG GGGACCCAGA AGACCGTGCC TTGCCCGGAA   60
GTCCTGCCTG TAGGCCTGAA GGACTTGCCC TCAACAACTA CCTGGTGATT TCAACAACTA  120
CCTACTTCAG CCCCTGGTG TGAGCAGCTT CTCAACATGA ACTACAGCCT CCACTTGGCC  180
TTCGTGTGTC TGAGTCTCTT CACTGAGAGG ATGTGCATCG AGGGGGAGTCA GTTCAACGTC  240
GAGGTCGGCA GAAGTGACGA GCTTTCCCTG CCTGGCCTTTG AGAACCCTCAC AGCAGGGGAT AT  300
AACAAATTTC TCAGGCCCAA TTTTGGTGGA GAACCCGTAC AGATAGCGCT GACTCTGGAC  360
ATTGCAAGTA TCTCTAGCAT TTCAGAGAGT AACATGGACT ACACAGCCAC CATATACCTC  420
CGACAGCGCT GGATGGACCA GCGGCTGGTG TTTGAAGGCA ACAGAGCTT CACTCTGGAT  480
GCCCGCCTCG TGGAGTTCCT CTGGGTGCCA GATACTTACA TTGTGGAGTC CAAGAAGTCC  540
TTCCTCCATG AAGTCACTGT GGGAAACAGG CTCATCCGCC TCTTCTCCAA TGGCACGGTC  600
CTGTATGCCC TCAGAATCAC GACAACTGTT GCATGTAACA TGGATCTGTC TAAATACCCC  660
ATGGACACAC AGACATGCAA GTTGCAGCTG GAAAGCTGGG GCTATGATGG AAATGATGTG  720
GAGTTCACCT GGCTGAGAGG GAACGACTCT GTGCGTGGAC TGGAACACCT GCCGGCTTGCT  780
CAGTACACCA TAGAGCGGTA TTTCACCTTA GTCACCAGAT CGCAGCAGGA GACAGGAAAT  840
TACACTAGAT TGGTCTTACA GTTTGAGCTT CGGAGGAATG TTCTGTATTT CATTTTGGAA  900
ACCTACGTTC CTTCCACTTT CCTGGTGGTG TTGTCCTGGG TTTCATTTTG GATCTCTCTC  960
GATTCAGTCC CTGCAAGAAC CTGCCATTGGA GTGACGACCG TGTTATCAAT GACCACACTG 1020
ATGATCGGGT CCCGCACTTC TCTTCCCAAC ACCAACTGCT TCATCAAGGC CATCGATGTG 1080
TACCTGGGGA TCTGCTTTAG CTTTGTGTTT GGGGCCTTGC TAGAATATGC AGTTGCTCAC 1140
TACAGTTCCT TACAGCAGAT GGCAGCCAAA GATAGGGGGA CAACAAAGGA AGTAGAAGAA 1200
GTCAGTATTA CTAATATCAT CAACAGCTCC ATCTCCAGCT TTAAACGGAA GATCAGCTTT 1260
GCCAGCATTG AAATTTCCAG CGACAACGTT GACTACAGTG ACTTGACAAT GAAAACCAGC 1320
GACAAGTTCA AGTTTGTCTT CCGAGAAAAG ATGGGCAGGA TTGTTGATTA TTTCACAATT 1380
CAAAACCCCA GTAATGTTGA TCACTATTCC AAACTACTGT TTCCTTTGAT TTTTATGCTA 1440
GCCAATGTAT TTTACTGGGC ATACTACATG TATTTTGAG TCAATGTTAA ATTTCTTGCA 1500
TGCCATAGGT CTTCAACAGG ACAAGATAAT GATGTAAATG GTATTTTAGG CCAAGTGTGC 1560
ACCCCACATCC AATGGTGCTA CAAGTGACTG AAATAATATT TGAGTCTTTC TGCTCAAAGA 1620
ATGAAGCTCC AACCATTGTT CTAAGCTGTG TAGAAGTCCT AGCATTATAG GATCTTGTAA 1680
TAGAAACATC AGTCCATTCC TCTTTCATCT TAATCAAGGA CATTCCCATG GAGCCCAAGA 1740
TTACAAATGT ACTCAGGGCT GTTTATTCGG TGGCTCCCTG GTTTGCATTT ACCTCATATA 1800
AAGAATGGGA AGGAGACCAT TGGGTAACCC TCAAGTGTCA GAAGTTGTTT CTAAAGTAAC 1860
TATACATGTT TTTTACTAAA TCTCTGCAGT GCTTATAAAA TACATTGTTG CCTATTTAGG 1920
GAGTAACATT TTCTAGTTTT TGTTTCTGGT TAAAATGAAA TATGGGCTTA TGTCAATTCA 1980
```

```
         TTGGAAGTCA ATGCACTAAC TCAATACCAA GATGAGTTTT TAAATAATGA ATATTATTTA   2040
         ATACCACAAC AGAATTATCC CCAATTTCCA ATAAGTCCTA TCATTGAAAA TTCAAATATA   2100
         AGTGAAGAAA AAATTAGTAG ATCAACAATC TAAACAAATC CCTCGGTTCT AAGATACAAT   2160
5        GGATTCCCCA TACTGGAAGG ACTCTGAGGC TTTATTCCCC CACTATGCAT ATCTTATCAT   2220
         TTTATTATTA TACACACATC CATCCTAAAC TATACTAAGA CCCTTTTCCC ATGCATGGAT   2280
         GGAAATGGAA GATTTTTTTG TAACTTGTTC TAGAAGTCTT AATATGGGCT GTTGCCATGA   2340
         AGGCTTGCAG AATTGAGTCC ATTTTCTAGC TGCCTTTATT CACATAGTGA TGGGGTACTA   2400
         AAAGTACTGG GTTGACTCAG AGAGTCGCTG TCATTCTGTC ATTGCTGCTA CTCTAACACT   2460
10       GAGCAACACT CTCCCAGTGG CAGATCCCCT GTATCATTCC AAGAGGAGCA TTCATCCCTT   2520
         TGCTCTAATG ATCAGGAATG ATGCTTATTA GAAAACAAAC TGCTTGACCC AGGAACAAGT   2580
         GGCTTAGCTT AAGTAAACTT GGCTTTGCTC AGATCCCTGA TCCTTCCAGC TGGTCTGCTC   2640
         TGAGTGGCTT ATCCCGCATG AGCAGGAGCG TGCTGGCCCT GAGTACTGAA CTTTCTGAGT   2700
         AACAATGAGA CACGTTACAG AACCTATGTT CAGGTTGCGG GTGAGCTGCC CTCTCCAAAT   2760
15       CCAGCCAGAG ATGCACATTC CTCGGCCAGT CTCAGCCAAC AGTACCAAAA GTGATTTTTG   2820
         AGTGTGCCAG GGTAAAGGCT TCCAGTTCAG CCTCAGTTAT TTTAGACAAT CTCGCCATCT   2880
         TTAATTTCTT AGCTTCCTGT TCTAATAAAT GCACGGCTTT ACCTTTCCTG TCAGAAATAA   2940
         ACCAAGGCTC TAAAAGATGA TTTCCCTTCT GTAACTCCCT AGAGCCACAG GTTCTCATTC   3000
         CTTTTCCCAT TATACTTCTC ACAATTCAGT TTCTATGAGT TTGATCACCT GATTTTTTTA   3060
20       ACAAAATATT TCTAACGGGA ATGGGTGGGA GTGCTGGTGA AAAGAGATGA AATGTGGTTG   3120
         TATGAGCCAA TCATATTTGT GATTTTTTAA AAAAAGTTTA AAAGGAAATA TCTGTTCTGA   3180
         AACCCCACTT AAGCATTGTT TTTATATAAA AACAATGATA AAGATGTGAA CTGTGAAATA   3240
         AATATACCAT ATTAGCTACC CACC
```

Seq ID NO: 54 Protein sequence
Protein Accession #: NP_055026

```
         1          11         21         31         41         51
         |          |          |          |          |          |
         MNYSLHLAFV CLSLFTERMC IQGSQFNVEV GRSDKLSLPG FENLTAGYNK FLRPNFGGEP    60
         VQIALTLDIA SISSISESNM DYTATIYLRQ RWMDQRLVFE GNKSFTLDAR LVEFLNVPDT   120
         YIVESKKSFL HEVTVGNRLI RLFSNGTVLY ALRITTTVAC NMDLSKYPMD TQTCKLQLES   180
         WGYDGNDVEF TWLRGNDSVR GLEHLRLAQY TIERYFLVT RSQQETGNYT RLVLQFELRR   240
         NVLYFILETY VPSTFLVVLS WVSFWISLDS VPARTCIGVT TVLSMTTLMI GSRTSLPNTN   300
         CFIXAIDVYL GICFSFVFGA LLEYAVAHYS SLQQMAAKDR GTTKEVEEVS ITNIINSSIS   360
         SFKRKISFAS IEISSDNVDY SDLTMKTSDK FKPVFREKMG RIVDYFTIQN PSNVDHYSKL   420
         LFPLIFMLAN VPYWAYYMYF
```

Seq ID NO: 55 DNA sequence
Nucleic Acid Accession #: XM_084007
Coding sequence: 138..2405

```
         1          11         21         31         41         51
         |          |          |          |          |          |
         CTCGTGCCGA ATTCGGCACG AGACCGCGTG TTCGCGCCTG GTAGAGATTT CTCGAAGACA    60
         CCAGTGGGCC CGTGTGGAAC CAAACCTGCG CGCGTGGCCG GGCCGTGGGA CAACGAGGCC   120
         GCGGAGACGA AGGCGCAATG GCGAGGGAAGT TATCTGTAAT CTTGATCCTG ACCTTGCCC   180
         TCTCTGTCAC AAATCCCCTT CATGAACTAA AAGCAGCTGC TTTCCCCCAG ACCACTGAGA   240
         AAATTAGTCC GAATTGGGAA TCTGGCATTA ATGTTGACTT GGCAATTTCC ACACGGCAAT   300
         ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAAATAA TTCTTTGTCA GTTGAAGGGT   360
         TCAGAAAATT ACTTCAAAAT ATAGGCATAG ATAAGATTAA AAGAATCCAT ATACACCATG   420
         ACCACGACCA TCACTCAGAC CACGAGCATC ACTCAGACCA TGAGCGTCAC TCAGACCATG   480
         AGCATCACTC AGACCACGAG CATCACTCTG ACCATGATCA TCACTCCCAC CATAATCATG   540
         CTGCTTCTGG TAAAAATAAG CGAAAAGCTC TTTGCCCAGA CCATGACTCA GATAGTTCAG   600
         GTAAAGATCC TAGAAACAGC CAGGGGAAAG GAGCTCACCG ACCAGAACAT GCCAGTGGTA   660
         GAAGGAATGT CAAGGACAGT GTTAGTGCTA GTGAAGTGAC CTCAACTGTG TACAACACTG   720
         TCTCTGAAGG AACTCACTTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAACTCTTCC   780
         CCAAAGATGT AAGCAGCTCC ACTCCACCCA GTGTCACATC AAAGAGCCGG GTGAGCCGGC   840
         TGGCTGGTAG GAAAACAAAT GAATCTGTGA GTGAGCCCCG AAAAGGCTTT ATGTATTCCA   900
         GAAACACAAA TGAAAATCCT CAGGAGTGTT TCAATGCATC AAAGCTACTG ACATCTCATG   960
         GCATGGGCAT CCAGGTTCCG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA  1020
         TCAACCAAAT TGATGCTAGA TCTTGTCTGA TTCATACAAG TGAAAAGAAG GCTGAAATCC  1080
         CTCCAAAGAC CTATTCATTA CAAATAGCCT GGGTTGGTGG TTTTATAGCC ATTTCCATCA  1140
         TCAGTTTCCT GTCTCTGCTG GGGGTTATCT TAGTGCCTCT CATGAATCGG GTGTTTTTCA  1200
         AATTTCTCCT GAGTTTCCTT GTGGCACTGG CCGTTGGGAC TTTGAGTGGT GATGCTTTTT  1260
         TACACCTTCT TCCACATTCT CATGCAAGTC ACCACCATAG TCATAGCCAT GAAGAACCAG  1320
         CAATGGAAAT GAAAAGAGGA CCACTTTTCA GTCATCTGTC TTCTCAAAAC ATAGAAGAAA  1380
         GTGCCTATTT TGATTCCACG TGGAAGGGTC TAACAGCTCT AGGAGGCCTG TATTTCATGT  1440
         TTCTTGTTGA ACATGTCCTC ACATTGATCA AACAATTTAA AGATAAGAAG AAAAAGAATC  1500
         AGAAGAAACC TGAAAATGAT GATGATGTGG AGATTAAGAA GCAGTTGTCC AAGTATGAAT  1560
         CTCAACTTTC AACAAATGAG GAGAAAGTAG ATACAGATGA TCGAACTGAA GGCTATTTAC  1620
         GAGCAGACTC ACAAGAGCCC TCCCACTTTG ATTCTCAGCA GCCTGCAGTC TTGGAAGAAG  1680
         AAGAGGTCAT GATAGCTCAT GCTCATCCAC AGGAAGTCTA CAATGAATAT GTACCCAGAG  1740
         GGTGCAAGAA TAAATGCCAT TCACATTTCC ACGATACACT CGGCCAGTCA GACGATCTCA  1800
         TTCACCACCA TCATGACTAC CATCATATTC TCCATCATCA CCACCACCAA AACCACCATC  1860
         CTCACAGTCA CAGCCACCGC TACTCTCGGG AGGAGCTGAA AGATGCCGGC GTCGCCACTT  1920
         TGGCCTGGAT GGTGATAATG GGTGATGGCC TGCACAATTT CAGCGGATGGC CTAGCAATTG  1980
         GTGCTGCTTT TACTGAAGGC TTATCAAGTG GTTTAAGTAC TTCTGTTGCT GTGTTCTGTC  2040
         ATGAGTTGCC TCATGAATTA GGTGACTTTG CTGTTCTACT AAAGGCTGGC ATGACCGTTA  2100
         AGCAGGCTGT CCTTTATAAT GCATTGTCAG CCATGCTGGC GTATCTTGGA ATGGCAACAG  2160
         GAATTTTCAT TGGTCATTAT GCTGAAAATG TTTCTATGTG GATATTTGCA CTTACTGCTG  2220
         GCTTATTCAT GTATGTTGCT CTGGTTGATA TGGTACCTGA AATGCTGCAC AATGATGCTA  2280
         GTGACCATGG ATGTAGCCGC TGGGGGTATT TCTTTTTACA GAATGCTGGG ATGCTTTTGG  2340
         GTTTTGGAAT TATGTTACTT ATTTCCATAT TTGAACATAA AATGTGTTTT CGTATAAATT  2400
         TCTAGTTAAG GTTTAAATGC TAGAGTAGCT TAAAAAGTTG TCATAGTTTC AGTAGGTCAT  2460
```

```
     AGGGAGATGA GTTTGTATGC TGTACTATGC AGCGTTTAAA GTTAGTGGGT TTTGTGATTT   2520
     TTGTATTGAA TATTGCTGTC TGTTACAAAG TCAGTTAAAG GTACGTTTTA ATATTTAAGT   2580
     TATTCTATCT TGGAGATAAA ATCTGTATGT GCAATTCACC GGTATTACCA GTTTATTATG   2640
     TAAACAAGAG ATTTGGCATG ACATGTTCTG TATGTTTCAG GGAAAAATGT CTTTAATGCT   2700
5    TTTTCAAGAA CTAACACAGT TATTCCTATA CTGGATTTTA GGTCTCTGAA GAACTGCTGG   2760
     TGTTTAGGAA TAAGAATGTG CATGAAGCCT AAAATACCAA GAAAGCTTAT ACTGAATTTA   2820
     AGCAAAGAAA TAAAGGAGAA AAGAGAAGAA TCTGAGAATT GGGGAGGCAT AGATTCTTAT   2880
     AAAAATCACA AAATTTGTTG TAAATTAGAG GGGAGAAATT TAGAATTAAG TATAAAAAGG   2940
     CAGAATTAGT ATAGAGTACA TTCATTAAAC ATTTTTGTCA GGATTATTTC CCGTAAAAAC   3000
10   GTAGTGAGCA CTCTCATATA CTAATTAGTG TACATTTAAC TTTGTATAAT ACAGAAATCT   3060
     AAATATATTT AATGAATTCA AGCAATATAC ACTTGACCAA GAAATTGGAA TTTCAAAATG   3120
     TTCGTGCGGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGTATCAC CAGACTGGGT   3180
     TATTGCCAAG TTATATATCA CCAAAAGCTG TATGACTGGA TGTTCTGGTT ACCTGGTTTA   3240
     CAAAATTATC AGAGTAGTAA AACTTTGATA TATATGAGGA TATTAAAACT ACACTAAGTA   3300
15   TCATTTGATT CGATTCAGAA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCATTGT   3360
     GAGCAATTGT CTTTATATAC GGTACTGTAG CCATACTAGG CCTGTCTGTG GCATTCTCTA   3420
     GATGTTTCTT TTTTACACAA TAAATTCCTT ATATCAGCTT G
```

Seq ID NO: 56 Protein sequence
Protein Accession #: XP_084007

```
     1         11        21        31        41        51
     |         |         |         |         |         |
     MARKLSVILI LTFALSVTNP LHELKAAAFP QTTEKISPNW ESGINVDLAI STRQYHLQQL   60
     FYRYGENNSL SVEGFRKLLQ NIGIDKIKRI HIHHDHDHHS DHEHHSDHER HSDHEHHSDH   120
     EHHSDHDHHS HHNHAASGKN KRKALCPDHD SDSSGKDPRN SQGKGAHRPE HASGRRNVKD   180
     SVSASEVTST VYNTVSEGTH FLETIETPRP GRLFPKDVSS STPPSVTSKS RVSRLAGRKT   240
     NESVSEPRKG FMYSRNTNEN PQECFNASKL LTSHGMGIQV PLNATEFNYL CPAIINQIDA   300
     RSCLIHTSRK KAEIPPKTYS LQIAWVGGFI AISIISFLSL LGVILVPLMN RVFFKFLLSP   360
     LVALAVGTLS GDAFLHLLPH SHASHHHSHS HEEPAMEMKR GPLFSHLSSQ NIEESAYFDS   420
     TWKGLTALGG LYFMFLVEHV LTLIKQFKDK KKKNQKKPEN DDDVEIKKQL SKYESQLSTN   480
     EEKVDTDDRT EGYLRADSQE PSHFDSQQPA VLEEEEVMIA HAHPQEVYNE YVPRGCKNKC   540
     HSHFHDTLGQ SDDLIHHHHD YEHILHHHHH QNHHPHSHSQ RYSREELKDA GVATLAWMVI   600
     MGDGLHNFSD GLAIGAAFTE GLSSGLSTSV AVFCHELFHE LGDFAVLLKA GMTVKQAVLY   660
     NALSAMLAYL GMATGIFIGH YAENVSMWIF ALTAGLFMYV ALVDMVPEML HNDASDHGCS   720
     RWGYFFLQNA GMLLGFGIML LISIFEHKIV FRINF
```

Seq ID NO: 57 DNA sequence
Nucleic Acid Accession #: NM_015419.1
Coding sequence: 1..8487

```
     1         11        21        31        41        51
     |         |         |         |         |         |
     ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT TTGGGGCCAT   60
     CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT ACGTCCCCAG CGAGGTCCAC   120
     TGCACGTTCC GATCCCTGGC TTCCGTGCCC GCTGGCATTG CTAGACACGT GGAAAGAATC   180
     AATTTGGGGT TTAATAGCAT ACAGGCCCTG TCAGAAACCT CATTTGCAGG ACTGACCAAG   240
     TTGGAGCTAC TTATGATTCA CGGCAATGAG ATCCCAAGCA TCCCCGATGG AGCTTTAAGA   300
     GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT GATCACAGGA   360
     CAGACCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA TTGACCACAA CAAGATCGAG   420
     TTTATCCACC CTCAAGCTTT CAACGGCTTA ACGTCTCTGA GGCTACTCCA TTTGAAGGA   480
     AATCTCCTCC ACCAGCTGCA CCCCAGCACC TTCTCCACGT TCACATTTTT GGATTATTTC   540
     AGACTCTCCA CCATAAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCCTGCC   600
     AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAA TCCGTGGACC   660
     TGCGATTGTG AGATGAGATG GTTTTTGGAA TGGGATGCAA AATCCAGAGG AATTCTGAAG   720
     TGTAAAAAGG ACAAAGCTTA TGAAGGCGGT CAGTTGTGTG CAATGTGCTT CAGTCCAAAG   780
     AAGTTGTACA AACATGAGAT ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA   840
     GAGTCCCCTC TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG   900
     GATGGTGAGA GCCAGCTCAT CCTGGAGAAA TTCCAACTGC CCAGTGGAG CATCTCTTTG   960
     AAATGACCG ACGAGCACGG GAACATGGTG AACTTGGTCT GTGACATCAA GAAACCAATG   1020
     GATGTGTACA AGATTCACTT GAACCAAACG GATCCTCCAG ATATTGACAT AAATGCAACA   1080
     GTTGCCTGG ACTTTGAGTG TCCAATGACC CGAGAAAACT ATGAAAAGCT ATGGAAATTG   1140
     ATAGCCATACT ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC   1200
     CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTTA CTACACAGGT   1260
     GTGAGAGCCC AGATTCTTGC AGAACCAGAA TGGGTCATCG AGCCATCCAT AGATATCCAG   1320
     CTGAACCGAC GTCAGAGTAC GGCCAAGAAG GTGCTACTTT CCTACTACAC CCAGTATTCT   1380
     CAAACAATAT CCACCAAAGA TACAAGGCAG GCTCGGGGCA GAAGCTGGGT AATGATTGAG   1440
     CCTAGTGGAG CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC   1500
     TGCAACGTGA AAGCTTCTGA GAGTCCATCT ATCTTCTGGA TGCTTCCAGA TGGCTCCATC   1560
     CTGAAAGCGC CCATGGGATGA CCCAGACAGC AAGTTCTCCA TTCTCAGCAG TGGCTGGCTG   1620
     AGGATCAAGT CCATGGAGCC ATCTGACTCA GGCTTGTACC AGTGCATTGC TCAAGTGAGG   1680
     GATGAAATGG ACGGCATGGT ATATAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGCCAGCC   1740
     GAGAAAGACA CAGTGACAAT GGCAAGAAC CCAGGGGAGT CGGTGACATT GCCTTGCAAT   1800
     GCTTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG GATAATTAAT   1860
     GATTTGGCTA ACACATCACA TGTATACATG TTGCCAAATG GAACTCTTTC CATCCCAAAG   1920
     GTCCAAGTCA GTGATAGTGG TTACTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCACAC   1980
     CATTTTACGG TGGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC   2040
     AGACGCCCAG GTGCAAAGGC TCTTTCCAGA GTCAGAGAAA ACATCGTGGA GGATGAAGGG   2100
     GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC TTCTGCATCC AAAGGACCAA   2160
     GAGGTGTTCC TCAAAACAAA GGATGATGCC ATCAATGGAG ACAAGAAAGC CAAGAAAGGG   2220
     AGAAGAAAGC TGAAACTCTG GAAGCATTCG GAAAAAGAAC CAGAGACCAA TGTTGCAGAA   2280
     GGTCGCAGAG TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG   2340
     GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGGAAAA ATCTCCCTAA GGGCACAGAA   2400
     GTACCCCCGT TGATTAAAAC CACAAGTCCT CCATCCTTGA GCCTAGAAGT CACACCACCT   2460
```

```
TTTCCTGCTG TTTCTCCCCC CTCAGCATCT CCTGTGCAGA CAGTAACCAG TGCTGAAGAA 2520
TCCTCAGCAG ATGTACCTCT ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCCTCA 2580
GCCAGCATGG GGCTAGAACA CAACCACAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA 2640
AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA GATAACTTCC 2700
ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA TATCTGAGCC TTATGAACCA 2760
TCTCCTACTC TGCACACATT AGACACAGTC TATGAAAAGC CCACCCATGA AGAGACGGCA 2820
ACAGAGGGTT GGTCTGCAGC AGATGTTGGA TCGTCACCAG AGCCCACATC CAGTGAGTAT 2880
GAGCCTCCAT TGGATGCTGT CTCCTTGGCT GAGTCTGGAC CCATGCAATA CTTTGACCCA 2940
GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC CTTTGCACAC 3000
CTTACTCCAA CCCCCACCAT CTGGGTTAAT GACTCCAGTA CATCACAGTT ATTTGAGGAT 3060
TCTACTATAG GGGAACCAGG TGTCCCAGGC CAATCACATC TACAAGGACT GACAGACAAC 3120
ATCCACCTTG TGAAAAGTAG TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG 3180
AAAGAGATGT CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC 3240
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA CTCCACACTG 3300
GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG AAACCACAGT TGGTACCCTC 3360
CTAGACAAAG ACACCACACA AGTAACAACA ACACCAAGGC AAAAGTTGC TCCGTCATCC 3420
ACCATGAGCA CTCACCCTTC TCGAAGGAGA CCCAACGGGA GAAGGAGATT ACGCCCCAAC 3480
AAATTCCGCC ACCGGCACAA GCAAACCCCA CCCAACTT TTGCCCCATC AGAGACTTTT 3540
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA GAGTTCTCTG 3600
GTTCCTACAG CTTGGGTGGA TAACACAGTT AATACCCCCA AACAGTTGGA AATGGAGAAG 3660
AATGCAGAAC CCACATCCAA GGGAACACCA CGGAGAAAAC ACGGGAAGAG GCCAAACAAA 3720
CATCGATATA CCCCTTCTAC AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT 3780
CCAGAAAATA AACATAGAAA CATTGTTACT CCCAGTTCAG AAACTATACT TTTGCCTAGA 3840
ACTGTTTCTC TGAAAACTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC AACCACCAGA 3900
AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC CAGTCACATA TAAACCCACA 3960
TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG TTGACAAACA TAAAAAGTGAC 4020
ATTTTAGTCA CTGGTGAATC AATTACTAAT GCCATACCAA CTTCTCGCTC CTTGGTCTCC 4080
ACTATGGGAG AATTTAAGGA AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG 4140
AATCCCTCAA GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT 4200
GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA TTTCACTTCC 4260
GAGTTTTTGT CCTCTTGAC AGTCTCCACA CCATTCACC AGGAAGAAGC TGGTTCTTCC 4320
ACAACTCTCT CAAGCATAAA AGTGGAGGTG GCTTCAAGTC AGGCAGAAAC CACCACCCTT 4380
GATCAAGATC ATCTTGAAAC CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT 4440
CACACCCCTA CTGCTGCCCG GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC 4500
ATGTCTTTGG GACAAACCAC CACCCACTAAG CCAGCACTTC CCAGTCCAAG AATATCTCAA 4560
GCATCTAGAG ATTCCAAGGA AAATGTTTTC TTGAATTATG TGGGGAATCC AGAAACAGAA 4620
GCAACCCCAG TCAACAATGA AGGAACACAG CATATGTCAG GGCCAAATGA ATTATCAACA 4680
CCCTCTTCCG ACCGGGATGC ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA 4740
TTTGGTAGTA GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT 4800
GCTTCTCATC AACTAACCAG GTCCCTGCC AAACCCATCC TACCACAGC ACAGTGAGG 4860
CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG TAACTTCCCA GTCACCTCGT 4920
CACTGGACCA ACAAACCGGA AATAACTACA TATCCTTCTG GGGCTTTGCC AGAGAACAAA 4980
CAGTTTACAA CTCCAAGATT ATCAAGTACA ACAAATTCCTC TCCCATTGCA CATGTCCAAA 5040
CCCAGCATTC CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTACTCCAAA 5100
GTGTTTGGAA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC TCCCAGTCCA 5160
AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA CCAACAAGAC TCTTTCTTTT 5220
CCACAGTTGG GAGTCACCCG GAGACCCCAG ATACCCACTT CTCCTGCCCC AGTAATGAGA 5280
GAGAGAAAG TTATTCCAGG TTCCTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG 5340
GACTTTGGCC CTCCGGCACC TCCGGTTGTTG CACACTCCGC AGACCACGGG ATCACCCTCA 5400
ACTAACTTAC AGAATATCCC TATGGTCTCT TCCAACCAGA GTTCTATCTC CTTTATAACA 5460
TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT CAAAGTTCTT TGCAGGAGGA 5520
CCTCCTGCAT CCAAATTCTG GTCTCTTGGG GAAAAGCCCC AAATCCTCAC CAAGTCCCCA 5580
CAGACTGTGT CGGTCACCGC TGAAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA 5640
CCAAAGCCTT TGGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC 5700
AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA GGTTCAAGTA 5760
CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC ACGGCCTGGA CAGGATGGTG 5820
GTCTTGCTTT CGGTCACCGT GCAGCAACCT CAAATCCTAG CCTCCCACTA CCAGGACGTC 5880
ACTGTCTACC TGGGAGACAC CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC 5940
CAAATTTCCT GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC ATGGAGAGC 6000
CGCATCACCC TGCACGAAAA CGGACCCTT TCCATCAAGG AGGCGTCCTT CTCAGACAGA 6060
GGCGTCTATA AGTGCGTGGC CAGCAATGCA GCCGGGGCGG ACAGCCTGGC CATCCGCCTG 6120
CACGTGGCGG CACTGCCCCC CGTTATCCAC CAGGAGAACG TGGAGAACAT CTCGCTGCCC 6180
CCGGGGCTCA GCATTCACAT TCACTGCACT GCCAAGGCTG CGCCCCTGCC CAGCGTGCGC 6240
TGGGTGCTCG GGGACGGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG GAACTTGTTT 6300
GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC CCAAGGACAG CGGGCGCTAT 6360
GAGTGCGTGG CCGCCAACCT GGTAGGCTCC GCGCGCAGGA CGGTGCAGCT GAACGTGCAG 6420
CGTGCAGCAG CCAACGCGCG CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC 6480
GGAGGAACCC TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG 6540
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG AATCAAGGTG 6600
TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA AAGATGCCGG AGATTACCTG 6660
TGCGTAGCTC GAAATAAGGT TGGTGATGAC TACGTGGTGC TCAAAGTGGA TGTGGTGATG 6720
AAACCAGCCG AGATTGAACA CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTAG 6780
CTGAAAGTGG ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCCTG GAGCCTCCCA 6840
GACGGGAGTC TGGTGAACTC CTTCATGCAG TCGGATGACA GCGGTGGACG CACCAAGCGC 6900
TATGTCGTCT TCAACAATGG GACACTCTAC TTTAACGAAG TGGGGATGAG GGAGGAAGGA 6960
GACTACACCT GCTTTGCTGA AAATCAGGTC GGGAAGGACG AGATGAGAGT CAGAGTCAAG 7020
GTGGTGACAG CGCCCGCCAC CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT 7080
GGAGACGTGG TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG 7140
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT ATACCAAGAT 7200
GGCACTCTCC TTATTCAGAA AGCCCAGCGG TCTGACAGCG GCAACTACAC CTGCCTGGTC 7260
AGGAACAGCG CGGGAGAGGA TAGGAAGACG GTGTGGATTC ACGTCAACGT CCAGCCACCC 7320
AAGATCAACG GTAACCCCAA CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT 7380
CGGAAACTGA TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT 7440
CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT CCATGGCAAC 7500
```

970

```
GGTTCCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG TCCAGCTGGT ATGCATGGCA   7560
CGCAACGAGG GAGGGGAGGC GAGGTTGATC GTGCAGCTCA CTGTCCTGGA GCCCATGGAG   7620
AAACCCATCT TCCACGACCC GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC   7680
AGCCTCAACT GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT   7740
GGCACCGATC TGCAGAGTCGC ACAGCAGCTG CAGGCGCTTCT ACCACAAGGC TGACGGCATG   7800
CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT ACCGCTGCGT GGCCCGCAAT   7860
GCCGCTGGCC ACACGGAGAG GCTGGTCTCC CTGAAGGTGG GACTGAAGCC AGAAGCAAAC   7920
AAGCAGTATC ATAACCTGGT CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC   7980
CCTCCCGGGG CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG   8040
GGCCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC GGTTCGTGAG   8100
GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG AGACGGAGTA CGGCCCTTCG   8160
GTCACCAGCA TCCCCGTGAT TGTGATCGCC TATCCTCCCC GGATCACCAG CGAGCCCACC   8220
CCGGTCATCT ACACCCXGCC CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT   8280
CCCAAAGCTG ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTCAG   8340
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCCAGGGAT CACTGACCAT CCAGCATGCC   8400
ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA ACATTCTCGG CAGTGACTCC   8460
AAAACAACTT ACATCCACGT CTTCTGAAAT GTGGATTCAA GAATGATTGC TTAGGAACTG   8520
ACAACAAAGC GGGGTTTGTA AGGGAAGCCA GGTGGGGAA TAGGAGCTCT TAAATAATGT   8580
GTCACAGTGC ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT   8640
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT GAGACACTTT   8700
CTTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG TCTGTGCTCT GACTGCAATT   8760
TTTCTTCTTT TGCAAATGCC ACTCGACTGC CTTCATAAGC GTCCATAGGA TATCTGAGGA   8820
ACATTCATCA AAAATAAGCC ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA   8880
TCACCTAGTT AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCCAGAT TGACAAGTCA   8940
TCTTTCAGTT ATTTCCTCTG TCACTTCAAA ACTCCAGCTT GCCCAATAAG GATTTAGAAC   9000
CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT TACATACATA CAGCTACCAT   9060
TTTATATGAA AAAGAAAAA CATTTCTTCC TGGAACTCAC TTTTTATATA ATGTTTTATA   9120
TATATATTTT TTCCTTTCAA ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT   9180
ATTAAAATTA ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA   9240
ATATAATTTT AAAAAAATTTC TCTCCAACCT CCTTCAATT CAGTCACCAC TGTTATATTA   9300
CCTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT AGATTTCCTT GTATGCAAAG   9360
TTTTTGTTGA AAGCTGTGCT CAGAGGAGGT GAGAGGAGAG GAAGGAGAAA ACTGCATCAT   9420
AACTTTACAG AATTGAATCT AGAGTCTTCC CCGAAAAGCC CAGAAACTTC TCTGCAGTAT   9480
CTGGCTTGTC CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC   9540
CCATGAATAA TACACGACCT GTTATTTCCA TGACTGCTTT ACTGTATTTT TAAGGTCAAT   9600
ATACTGTACA TTTGATAATA AAATAATATT CTCCCAAAAA AAAAA
```

Seq ID NO: 58 Protein sequence
Protein Accession #: NP_056234.1

```
     1         11        21        31        41        51
     |         |         |         |         |         |
    MPKRAHWGAL SVVLILLWGH PRVALACPHP CACYVPSEVH CTFRSLASVP AGIARHVERI    60
    NLGFNSIQAL SETSPAGLTK LELLMIHGNE IPSIPDGALR DLSSLQVFKF SYNRLRVITG   120
    QTLQGLSNLM RLHIDHNKIE FIHPQAFNGL TSLRLLHLEG NLLHQLHPST FSTFTFLDYF   180
    RLSTIRHLYL AENMVRTLPA SMLRNMPLLE NLYLQGNPWT CDCEMRWFLE WDAKSRGILK   240
    CKKDKAYEGG QLCAMCFSPK KLYKHEIHKL KDMTCLKPSI ESPLRQNRSR SIEEEQEQEE   300
    DGGSQLILEK FQLPQWSISL NMTDEHGNMV NLVCDIKKPM DVYKIHLNQT DPPDIDINAT   360
    VALDFRCPMT RENYEKLWKL IAYYSEVPVK LHRELMLSKD PRVSYQYRQD ADEEALYYTG   420
    VRAQILAEPE WVMQPSIDIQ LNRRQSTAKK VLLSYYTQYS QTISTKDTRQ ARGRSWVMIE   480
    PSGAVQRDQT VLEGGPCQLS CNVKASESPS IFWVLPDGSI LKAPMDDPDS KFSILSSGWL   540
    RIKSMEPSDS GLYQCIAQVR DEMDRMVYRV LVQSPSTQPA EKDTVTIGKN PGESVTLPCN   600
    ALAIPEAHLS WILPNRRIIN DLANTSHVYM LPNGTLSIPK VQVSDSGYYR CVAVNQQGAD   660
    HFTVGITVTK KGSGLPSKRG RRPGAKALSR VREDIVEDEG GSGMEDEENT SRRLLHPKDQ   720
    EVFLKTKIDA INGDKKAKKG RRRLKLWKHS EKEPETNVAE GRRVPESRRR INMANKQINP   780
    ERWADILAKV EGKNLPKGTE VPPLIKTTSP PSLSLEVTPP FPAVSPPSAS PVQTVTSAEE   840
    SSADVPLLGE EEHVLGTISS ASMGLEHNHN GVILVEPEVT STPLEEVVDD LSEKTEEITS   900
    TBGDLKGTAA FTLISEPYEP SPTLHTLDTV YEKPTHEETA TEGWSAARDVG SSPEPTSSEY   960
    EFPLDAVSLA ESEPMQYFDP DLETKSQPDR DKMKEDTPAH LTPTPTIWVN DSSTSQLFED  1020
    STIGEPGVPG QSHLQGLTDN IHLVXSSLST QDTLLIKKGM KEMSQTLQGG NMLEGDPTHS  1080
    RSSESESEQES KSITLPDSTL GIMSSMSFVK KPAETTVGTL LDKDTTTVTT TPRQKVAPSS  1140
    TMSTHPSRRR PNGRRRLRPN KFRHRHKQTP PTTFPAPSETF STQPTQAPDI KISSQVESSL  1200
    VPTANVDNTV NTPKQLEMEK NAEPTSKGTP RRKHGKRPNK HRYTPSTVSS RASGSKPSPS  1260
    PENKHRNIVT PSSETILLPR TVSLKTEGPY DSLDYMTTTR KIYSSYPKVQ ETLPVTYKPT  1320
    SDGKEIKDTV ATNVDKEKSD ILVTGESIIN AIPTSRSLVS TMGEFKEESS PVGFPGTPTW  1380
    NPSRTAQPGR LQTDIPVTTS GENLTDPPLL KELEDVDPTS EFLSSLTVST PFHQEEAGSS  1440
    TTLSSIKVEV ASSQAETTTL DQDHLETTVA ILLSETRPQN HTPTAARMKE PASSSPSTIL  1500
    MSLGQTTTTK FALPSFRISQ ASRDSKENVF LNYVGNPETE ATPVNNEGTQ HMSGPNELST  1560
    PSSDRDAFNL STKLELEKQV FGSRSLPRGP DSQRQDGRVH ASHQLTEVPA KPILPTATVR  1620
    LPEMSTQSAS RYFVTSQBPR HWTNKPEITT YPSGALPENK QFTTPRLSST TIPLPLHMSK  1680
    PSIPSKFTDR RTDQFNGYSK VFGNNNIPEA RNPVGKPSP RIPHYSNGRL PPFTNKTLSF  1740
    PQLGVTRRPQ IPTSPAPVMR ERKVIPGSYN RIHSHSTFHL DFGPPAPPLL HTPQTTGSPS  1800
    TNLQNIPMVS STQSSISFIT SSVQSSGSFH QSSSKFFAGG PPASKFWSLG EKPQILTKSP  1860
    QTVSVTAETD TVFPCEATGK PKPFVTNTKV STGALMTPNT RIQRFEVLKN GTLVIRKVQV  1920
    QDRGQYMCTA SNLHGLDRMV VLLSVTVQQP QILASHYQDV TVYLGDTIAM ECLAKGTPAP  1980
    QISWIFPDRR VWQTVSPVES RITLHENRTL SIKEASFSDR GVYKCVASNA AGADSLAIRL  2040
    HVAALPPVIH QEKLENISLP PGLSIHIHCT AKAAPLPSVR WVLGDGTQIR PSQFLEHNLF  2100
    VFPNGTLYIR NLAPKDSGRY ECVAANLVGS ARRTVQLNVQ RAAANARITG TSPRRTDVRY  2160
    GGTLKLDCSA SGDPWPRILW RLPSKRMIDA LFSFDSRIKV PANGTLVVKS VIDKDAGDYL  2220
    CVARNXVGDD YVVLKVDVVM KPAKIEHKEE NDHKVFYGGD LKVDCVATGL PNPEISWSLP  2280
    DGSLVNSFMQ SDDSGGRTKR YVVFNNGTLY PNEVGMRREG DYTCFAENQV GKDEMRVRVK  2340
    VVTAPATIRN KTYLAVQPFY GDVVTVACEA KGEPMPKVTW LSPTNKVIPT SSEKYQIYQD  2400
    GTLLIQKAQR SDSGNYTCLV RNSAGEDRKT VWIHVNVQPP KINGNPNPIT TVREIAAGGS  2460
    RKLIDCKABG IPTPRVLWAP PEGVVLPAPY YGNRITVHGN GSLDIRSLRK SDSVQLVCMA  2520
```

```
              RNEGGEARLI VQLTVLEPME KPIFHDPISE KITAMAGHTI SLNCSAAGTP TPSLVWVLPN 2580
              GTDLQSGQQL QRPYHKADGM LHISGLSSVD AGAYRCVARN AAGHTERLVS LKVGLKPEAN 2640
              KQYHNLVSII NGETLKLPCT PPGAGQGRFS WTLPNGMHLE GPQTLGRVSL LDNGTLTVRE 2700
              ASVFDRGTYV CRMETSYGPS VTSIPVIVIA YPPRITSEPT PVIYTRPGNT VKLNCMAMGI 2760
              PKADITWELP DKSHLKAGVQ ARLYGNRFLH PQGSLTIQHA TQRDAGFYKC MAKNILGSDS 2820
              KTTYIHVP
```

```
       Seq ID NO: 59 DNA sequence
       Nucleic Acid Accession #: Eos sequence
       Coding sequence: 1..5001
```

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       ATGCCAGGCA CAAAACTAAC CCGAACAGGC GCCCCAGCAG ACTACAGAGT GATATTGAAG   60
       ACCTCTCAAG AGGACGAATT GGATGTACCT GACGACATCA GCGTCCGGGT TATGTCATCT  120
       CAGTCTGTGC TTGTGTCCTG GGTGGATCCT GTTCTGGAAA AACAGAAGAA AGTTGTTGCA  180
       TCAAGACAGT ACACCGTGCG CTATCGAGAG AAGGGGGAAT TGGCCAGGTG GGATTATAAG  240
       CAGATCGCTA ACAGGCGTGT GCTGATTGAG AACCTGATTC CAGACACTGT GTATGAATTT  300
       GCAGTCCGTA TTTCACGGGG TGAAAGAGAT GGCAAATGGA GTACGTCAGT CTTCCAAAGA  360
       ACACCAGAAT CTGCCCCTAC CACAGCTCCT GAAAACTTGA ACGTCTGGCC AGTCAATGGC  420
       AAACCTACAG TTGTCGCTGC ATCTTGGGAT GCGCTACCAG AGACTGAGGG GAAAGTGAAA  480
       GTCTGTCTGC TGGACACAGG ACTGTTTTCA GTTTCCTCCT TCCAACCATC TGCCAAATCA  540
       TTTCAGAATA CATTCTTTCA TACGCCCCGG CTCTCAAACC ATTTGGAGCA AAGTCCCTCA  600
       CCTATCCTGG AGACACTACT TCTGCCCTGG TGGATGGTCT GCAGCCTGGG GAACGCTATC  660
       TTTTCAAAAT CCGGGCCACA AACAGGAGAG GCCTGGGACC TCACTCCAAA GCCTTCATTG  720
       TGGCTATGCC AACAAGAATG CAGCTGTACC CAGAAGGATT TCAGTTGTCT AGCTTACCTG  780
       ATCGATATCC AAACCAAACA AGTTAATAAA GATCCACAAC TGGAAGGGAG TGTTTTTGGA  840
       CCATGTTTTC TTTCTACTT  CCTCACATTT ATGCTGGATA TTGGCGGCTT TTCCTTCATT  900
       ATGTGCTATG AAGACCCANN TGTTTCTTCT TTGACAGGCA ATTCTTTAAA ATCTGTTGCA  960
       GCCAGTAAGG CGGATGTTCA GCAGAACACG GAGGACAATG GGAAACCCGA AAAACCTGAG 1020
       CCTTCCTCAC CTTCTCCCAG AGCTCCAGCT TCCTCCCAAC ACCCCTCTGT GCCTGGTTCT 1080
       CCCCAAGGGA GAAATGCCAA GGACCTTCTT CTTGACTTGA AGAACAAAAT ATTGGCTAAT 1140
       GGTGGGGCGC CCGAAAAACC CCAGCTTCGC GCCAAGAAGG CAGAGGAGCT GGATCTTCAG 1200
       TCGACAGAAA TCACTGGGGA GGAGGAGCTG GGTTCCGGGG AGGACTCGCC CATGTCACCC 1260
       TCAGACACCC AAGACCAGAA ACGGACCCTG AGGCCGCCAA GTAGACACGG CCACTCGGTG 1320
       GTTGCTCCCG GCAGGACTGC AGTGAGGGCC CGGATGCCAG CGCTGCCCCG AAGGGAAGGC 1380
       GTAGATAAGC CTGGCTTTTC CCTGGCCACG CAGCCCCGCC CAGGGGCGCC CCCCTCGGCT 1440
       TCGGCCTCTC CTGCCCACCA CGCGTCCACC CAGGGCACCT CTCATCGTCC TTCCCTGCCT 1500
       GCCAGCTTGA ATGACAACGA CTTGGTGGAC TCAGACGAAG ATGAGCGCGC TGTGGGCTCC 1560
       CTCCACCCCA AGGGCGCCTT CGCCCAGCCC CGGCCAGCCC TGTCCCCCAG CGGCCAGTCC 1620
       CGGTCCAGCG TTCTCCGCGGA CAGAAGCTCT GTGCACCCCG GCGCAAAGCC AGCCTCGCCG 1680
       GCGCGGAGGA CCCCCCATTC AGGGGCCGCA GAGGAAGATT CCAGTGCCTC AGCCCCACCC 1740
       TCAAGACTTT CTCCACCCCA TGGGGGATCA TCTCGGCTGC TGCCCACCCA GCCACACCTG 1800
       AGCTCTCCAC TTTCCAAGGG CGGGAAGGAT GGTGAGGAGA CCCCAGCCAC CAACTCCAAT 1860
       GCGGCCATCAC GGTCCACCAT GTCCTCCTCC GTCTCTTCTC ATCTCTCGTC CAGGACGCAG 1920
       GTCTCTGAGG GAGCGGAGGC TTCTGATGGT GAAAGCCACG GTGACGGCGA TAGGGAAGAC 1980
       GGCGGAAGGC AGGCGGAGGC CACGGCCCAG ACGCTGCGGG CCGGCCTGC  CTCTGGAACAC 2040
       TTCCATTTGC TCAGACACAA ACCCTTTGCT GCCAACGGGA GGTCTCCAAG CAGGTTCAGC 2100
       ATTGGGCGGG GACCTCGGCT GCAGCCCTCC AGCTCCCAC  AGTCGACTGT GCCCTCCCGA 2160
       GCCCACCCCA GGGTTCCCTC TCACTCTGAT TCCCACCCTA AGCTTAGCTC AGGTATCCAT 2220
       GGAGACGAGG AGGATGAGAA GCCGCTTCCT GCCACCGTTG TCAATGACCA CGTGCCTTCC 2280
       TCCTCCAGGC AGCCCATCTC CCGGGGCTGG GAGGACTTAA GGAGAAGCCC GCAGAGAGGG 2340
       GCCAGCCTGC ATCGGAAGGA ACCCATCCCA GAGAACCCCA AATCCACAGG GGCAGATACA 2400
       CATCCTCAGG GCAAGTACTC CTCCCTGGCC TCCAAGGCTC AGGATGTTCA ACAGAGCACA 2460
       GACGCGGACA CGGAGGGTCA TTCTCCCAAA GCACAGCCAG GGTCCACAGA CCGGCCACGCG 2520
       TCCCCTGCTC GTCCTCCCGC AGCACGGGTCA CAGCAGCATC CCAGTGTTCC CAGAAGGATG 2580
       ACACCCGGCC GGGCCCCAGA ACAGCAGCCC CCTCCTCCCG TCGCCACGTC CCAGCACCAC 2640
       CCGGGACCCC AGAGCAGAGA CGCGGGTCGG TCACCTTCCC AGCCCAGGCT CTCACTGACC 2700
       CAGGCCGGGC GGCCCCGCCC CACGTCGCAG GGCGCGCTCCC ACTCCTCCTC GGACCCTTAC 2760
       ACGGCGAGC  CCAGAGGGAT GCTCCCCACG GCCCTCCAGA ACCAGGACGA GGATGCCCAG 2820
       GGCAGCTACG ACGACGACAG CACAGAAGTC GAGGCCCAGG ATGTGCGGGC CCCGCGCGCAC 2880
       GCCGCGCGCG CCAAGGAGGC AGCTGCGTCC CTTCCCAAGC ACCAGCAGGT GGAGTCTCCC 2940
       ACAGGCGCAG GGGCAGGTGG CGACCACAGG TCCCAGCGCG CTCCCCGCGCCC 3000
       AGGCCCAGCC GACCCGGCGG CCCCCAGTCC CGCGCCCGGG TCCCCAGCAG GGCAGCGCCG 3060
       GGGAAGTCGG AGCCTCCTTC CAAGCGGCCC CTGTCCTCCA AGTCCCAGCA GTCGGTCTCA 3120
       GCCGAGGACG AGGAGGAGGA GGACGCGGGG TTTTTTAAGG GCGGGAAAGA AGACCTTCTG 3180
       TCTTCCTCTG TGCCAAAGTG GCCCTCTTCC TCCACTCCCA GGGCGGGCAA AGACGCCGAT 3240
       GGGAGCCTCG CCAAGGAAGA GAGGGGAGCCT GCCATCGCGC TTGCCCCTCG CGGAGGGGAGC 3300
       CTGGCTCCTG TGAAGCGACC TCTCCCCCCA CCTCCAGGCA GCTCCCCCAG GGCCTCCCGAC 3360
       GTCCCTTCCC GACCGCCGCC TCGCAGCGCT GCCACCGTGA GCCCCGTCGC GGGCACCCAC 3420
       CCCTGGCCGC GGTACACCAC GCGCGCCCCV CCTGGCCACT TCTCCACCAC CCCGATGCTG 3480
       TCCTTGCGCC AGAGGATGAT GCATGCCAGA TTCCGTAACC CTCTCTCCCG ACAGCCTGCC 3540
       AGACCCTCTT ACAGACAAGG TTATAATGGC AGACCAAATG TAGAAGGGAA AGTCCTTCCT 3600
       GGTAGTAATG GAAAACCGAA TGGACAGAGA ATTATCAATG GCCCTCAAGG AACAAAGTGG 3660
       GTTGTGGACC TTGATCGTGG GTTAGTATTG AATGCAGAAG GAAGGTACCT CCAAGATTCA 3720
       CATGGAAATC CTCTTCGGAT TAAACTAGGA GGAGATGGTC GAACCATTGT AGATCTGGAA 3780
       GGGACCCCCG TGGTGAGTCC TGACGGCCTC CCACTCTTTG GGCAGGGGCG ACATGGCACA 3840
       CCTCTGGCCA ATGCCCAAGA TAAGCCAATT TTGAGTCTTG GAGGGAAAGCC GCTGGTGGGC 3900
       TTGGAGGTCA TCAAAAAAAC CACCCATCCC CCTACCACTA CCAGTCAGCC CACCACTACT 3960
       ACGACGCCCC TGCCTACCAC TACAACCCCG AGGCCCACCA CTGCCACCAC CATGCAGCCC 4020
       ACCACTACTA CGACGCCCCT GCCTACCACT ACACCGAGGC CCACCACTGC CACCACCCGC 4080
       CGCACGACCA CCAGGCGTCC AACAACCACA GTCCGAACCA CTACGCGGAC AACCACCACC 4140
       ACCACCCCCA AACCCACCAC TCCCATCCCC ACCTGTCCCC CTGGGACCTT GGAACGGCAC 4200
       GACGATGATG GCAACCTGAT AATGAGCTCC AATGGGATCC CAGAGTGCTA CGCTGAAGAA 4260
```

```
GATGAGTTCT CAGGCTTGGA GACTGACACT GCAGTACCTA CGGAAGAGGC CTACGTTATA  4320
TATGATGAAG ATTATGAATT TGAGACGTCA AGGCCACCAA CCACCACTGA GCCTTCGACC  4380
ACTGCTACCA CACCGAGGGT GATCCCAGAG GAAGGCGCCA TCAGTTCCTT TCCTGAAGAA  4440
GAATTTGATC TGGCTGGAAG GAAACGATTT GTTGCTCCTT ACGTGACGTA CCTAAATAAA  4500
GACCCATCAG CCCCGTGCTC TCTGACTGAT GCACTGGATC ACTTCCAAGT GGACAGCCTG  4560
GATGAAATCA TCCCCAATGA CCTGAAGAAG AGTGATCTGC CTCCCCAGCA TGCTCCCCGC  4620
AACATCACCG TGGTGGCCGT GGAAGGTTGC CACTCATTTG TCATTGTGGA TTGGGACAAA  4680
GCCACCCCAG GAGATTTGGT CACAGGTTAT TTGGTTTACA GTGCATCCTA TGAAGATTTC  4740
ATCAGGAACA AGTTTTCCAC TCAAGCTTCA TCAGTAACTC ACTTGCCCAT TGAGAACCTA  4800
AAGCCCAACA CGAGGTATTA TTTTAAAGTG CAAGCACAAA ATCCTCATGG CTACGGACCT  4860
ATCAGCCCTT CGGTCTCATT TGTCACCGAA TCAGATAATC CTCTGCTTGT TGTGAGGCCC  4920
CCAGGCGGTG AGCTATCTGG ATCCCATTCG CTTTCAAACA TGATCCCAGC TACACGGACT  4980
GCCATGGACG GCAATATGTG AAGCGCACGT GGTATCGAAA GTTCGTGGGA GTTGTTCTTT  5040
GTAATTCACT GAGGTATAAA ATCTACCTCA GTGACAACCT GAAAGATACA TTCTACAGCA  5100
TTGGAGACAG CTGGGGAAGA GGTGAAGACC ATTGCCAATT TGTGGATTCA CACCTTGATG  5160
GAAGAACAGG GCCTCAGTCC TATGTAGAAG CCCTCCCTAC TATTCAAGGC TACTATCGCC  5220
AGTATCGTCA GGAGCCTGTC AGGTTTGGGA ACATCGGCTT CGGAACCCCC TACTACTATG  5280
TGGGCTGGTA CGAGTGTGGG GTCTCCATCC CTGGAAAGTG GTAATCACAG GACCGTCATG  5340
CTGCAAGCTT GCCCTGCCCA GCCCCACCAA CTAAGTCGCA CTAGGGGCTG TGAGCAAAGA  5400
CAGCCAGCAT GCTCAGCCCC GCTGCCCTAG GTGCCAGGAA GGTCACAGAT GGACACTGGC  5460
CATTCTGGTC ATCTCAGTCT GGAACTCAGT CCCACTTCTT GGCCTGGACA ATGAACAGGA  5520
TTCAGTTTTG CTGTTAACTT TGCTTCTCTA CTTTTTTTTG TTTGTTTGTA ATAGCACATC  5580
CCAGAGACAT CAGAAACCAG CAACTGATTC AGTGTGATTT CCCAGACTTT TTAGGCATGA  5640
AATTCGGACA CTTCAGTATT TCCAGGAATA GCATATGCAC GCTGTTCTTG CTTCATGGAA  5700
TGCTACATGC TTTCTGTTTT TCTCATTTTG GATTTCTCCA AAACTAACTG AATTTAAGCT  5760
TCAGGTCCCT TTGTATGCAG TAGAAAGGAA TTATTAAAAA CACCACCAAA GAAAATAAAT  5820
ATATCCTACT TGAAATTTAC TCTATGGACT TACCCACTGC TAGAATAAAT GTATCAAATC  5880
TTATTTGTAA ATTCTCAATT TTGATATATA TATGTATATA TGCATATACA TATCCACACT  5940
TGTCTGCAAG AATATTGATT AAAATTGCTA AATTTGTACT TGTTCACCAA AAAAAAAAAA  6000
AAAAAAA
```

Seq ID NO: 60 <u>Protein sequence</u>
Protein Accession #: Eos sequence

```
1          11         21         31         41         51
|          |          |          |          |          |
MFGTKLTRTG APADYRVILK TSQEDELDVP DDISVRVMSS QSVLVSWVDP VLEKQKKVVA   60
SRQYTVRYRE KGELARWDYK QIANRRVLIE NLIPDTVYEF AVRISQGERD GKWSTSVFQR  120
TPESAPTTAP ENLNVWPVNG KPTVVAASWD ALPETEGKVK VCLLDTGLFS VSSFQPSAKS  180
FQNTFFHTPR LSNHLEQSPS PILETLLLPW WMVCSLANAI FSKSGFQTGE AWDLTPKPSL  240
SLCQQECSCT QKDFSCLAYL IDIQTKQVNK DPQLEGSVPG PCPLFYFLTF MLDIGGFSPI  300
MCYEDPVSSL TGNSLKSVAA SKADVQQNTE DNGKPEKPEP SSPSERAPAS SQHPSVPASP  360
QGRNAKDLLL DLKNKILANG GAPRKPQLRA KKAEELDLQS TEITGEERLG SREDSPMSPS  420
DTQDQKRTLR PPSRHGHSVV APGRTAVRAR MPALPRREGV DKPGFSLATQ PRPGAPPSAS  480
ASPAHHASTQ GTSHRPSLPA SLNDNDLVDS DEDERAVGSL HPKGAFAQPR PALSPSRQSP  540
SSVLRDRSSV HPGAKPASPA RRTPHSGAAE EDSSASAPPS RLSPPHGGSS RLLPTQPHLS  600
SPLSKGGKDG EDAPATNSNA PSRSTMSSSV SSHLSSRTQV SEGAEASDGE SEGDGDREDG  660
GRQAEATAQT LRARPASGHF HLLRHKPFAA NGRSPSRFSI GEGPRLQPSS SPQSTVPSRA  720
HPRVPSHSDS HPKLSSGIHG DEEDEKPLPA TVVNDHVPSS SRQPISRGWE DLRRSPQRGA  780
SLHRKEPIPE NPKSTGADTH PQGKYSSLAS KAQDVQQSTD ADTEGHSPKA QPGSTDRHAS  840
PARPPAARSQ QHPSVPRRMT PGRAPEQQPP PPVATSQHHP GPQSRDAGRS PSQPRLSLTQ  900
AGRPRPTSQG RSHSSSDPYT ASSRGMLPTA LQNQDEDAQG SYDDDSTEVE AQDVRAPAHA  960
ARAKEAAASL PKHQQVESPT GAGAGGDERS QRGHAASPAR PSRPGGPQSR ARVPSRAAPG 1020
KSEPPSKRPL SSKSQQSVSA EDEEEEDAGF PKGGKEDLLS SSVPKWPSSS TPRGGKDADG 1080
SLAKEERREPA IALAPRGGSL APVKRPLPPP PGSSPRASEV PSRPPPRSAA TVSPVAGTHP 1140
WPRYTTRAPP GEFSTTPMLS LRQRMMEARF RNPLSRQPAR PSYRQGYNGR PNVEGKVLPG 1200
SNGKPNGQRI INGPQGTEHV VDLDRGLVLN AEGRYLQDSH GNPLRIKLGG DGRTIVDLEG 1260
TPVVSPDGLP LFGQGRHGTP LANAQDKPIL SLGGKPLVGL EVIKKTTHPP TTTMQPTTTT 1320
TPLPTTTTPR PTTATTMQPT TTTTPLPTTT PRPTTATTRR PTTTTV RTTTTTTTTT 1380
TPKPTTPIPT CPPGTLERHD DDGNLIMSSN GIPECYAEED EFSGLETDTA VPTERAYVIY 1440
DEDYEFETSR PPTTTEPSTT ATTPKVIPEE GAISSFPEEE FDLAGRKRFV APYVTYLNKD 1500
PSAPCSLTDA LDHFQVDSLD EIIPNDLKKS DLPPQHAPRN ITVVAVEGCH SFVIVDWDKA 1560
TPGDLVTGYL VYSASYEDPI RNKFSTQASS VTHLPIENLK PNTRYYFKVQ AQNPHGYGPI 1620
SPSVSFVIES DNPLLVVRPP GGELSGSHSL SNMIPATRTA MDGNM
```

Seq ID NO: 61 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_022743
Coding sequence: 128..1237

```
1          11         21         31         41         51
|          |          |          |          |          |
GTGGATTTTA GAGATACCTC CCCTCCTTCT GCTCAGCTGC CTTGCAGTAA TTAAACTCTT   60
TCTCTGCTGC AACACCCCTA CTGTTCTCCG TGTATTGGCT TTTCTGGGCA GCAGGAAGGA  120
AAAGCTGATG CGATGCTCTC AGTGCCGCGT CGCCAAATAC TGTAGTGCTA AGTGTCAGAA  180
AAAAGCTTGG CCAGACCAGA AGCGGGAATG CAAATGCCTT AAAAGCTGCA AACCCAGATA  240
TCCTCCAGAC TCCGTTCGAC TTCTTGGCAG AGTTGTCTTC AAACTTATGG ATGGAGCACC  300
TTCAGAATCA GAGAAGCTTT ACTCATTTTA TGATCTGGAG TCAAATATTA ACAAACTGAC  360
TGAAGATAAG AAAGAGGGCC TCAGGCAACT CGTAATGACA TTTCAACATT TCATGAGAGA  420
AGAAATACAG GATGCCTCTC AGCTGCCACC TGCCTTTGAC CTTTTTGAAG CCTTTGCAAA  480
AGTGATCTGC AACTCTTTCA CCATCTGTAA TGCGGAGATG CAGGAAGTTG GTGGTTGGCCT  540
ATATCCCAGT ATCTCTTTGC TCAATCACAG CTGTGACCCC AACTGTTCGA TTGTGTTCAA  600
TGGGCCCCAC CTCTTACTGC GAGCAGTCCG AGACATCGAG GTGGGAGAGG AGCTCACCAT  660
CTGCTACCTG GATATGCTGA TGACCAGTGA GGAGCGCCGG AAGCAGCTGA GGGACCAGTA  720
CTGCTTTGAA TGTGACTGTT TCCGTTGCCA AACCCAGGAC AAGGATGCTG ATATGCTAAC  780
```

```
       TGGTGATGAG CAAGTATGGA AGGAAGTTCA AGAATCCCTG AAAAAAATTG AAGAACTGAA   840
       GGCACACTGG AAGTGGGAGC AGGTTCTGGC CATGTGCCAG GCGATCATAA GCAGCAATTC   900
       TGAACGGCTT CCCGATATCA ACATCTACCA GCTGAAGGTG CTCGACTGCG CCATGGGATGC  960
  5    CTGCATCAAC CTCGGCCTGT TGGAGGAAGC CTTGTTCTAT GGTACTCGGA CCATGGGAGCC  1020
       ATACAGGATT TTTTTCCCAG GAAGCCATCC CGTCAGAGGA GTTCAAGTGA TGAAAGTTGG  1080
       CAAACTGCAG CTACATCAAG GCATGTTCC CCAAGCAATG AAGAATCTGA GACTGGCTTT    1140
       TGATATTATG AGAGTGACAC ATGGCAGAGA ACACAGCCTG ATTGAAGATT TGATTCTACT   1200
       TTTAGAAGAA TGCGACGCCA ACATCAGAGC ATCCTAAGGG AACGCAGTCA GAGGGAAATA   1260
       CGGCGTGTGT CTTTGTTGAA TGCCTTATTG AGGTCACACA CTCTATGCTT TGTTAGCTGT   1320
  10   GTGAACCTCT CTTATTGGAA ATTCTGTTCC GTGTTTGTGT AGGTAAATAA AGGCAGACAT   1380
       GGTTTGCAAA CCACAAGAAT CATTAGTTGT AGAGAAGCAC GATTATAATA AATTCAAAAC   1440
       ATTTGGTTGA GGATGCCAAA AAAAAAAAAA AAAAAAA
```

Seq ID NO: 62 Protein sequence
Protein Accession #: NP_073580

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       MRCSQCRVAK YCSAKCQKKA WPDHKRECKC LKSCKPRYPP DSVRLLGRVV FKLMDGAPSE   60
       SEKLYSFYDL ESNINKLTED KKEGLRQLVM TFQHFMREEI QDASQLPPAF DLFEAPAKVI   120
       CNSFTICNAE MQEVGVGLYP SISLLNHSCD PNCSIVFNGP HLLLRAVRDI EVGEELTICY   180
       LDMLMTSEER RKQLRDQYCF ECDCFRCQTQ DKDADMLTGD EQVWKEVQES LKKIEELKAH   240
       WKWEQVLAMC QAIISSNSER LPDINIYQLK VLDCAMDACI NLGLLEEALF YGTRTMEPYR   300
       IFFPGSHPVR GVQVMKVGKL QLHQGMFPQA MKNLRLAFDI MRVTHGREHS LIEDLILLLE   360
       ECDANIRAS
```

Seq ID NO: 63 DNA sequence
Nucleic Acid Accession #: NM_003014.2
Coding sequence: 238..648

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       GGCGGGTTCG CGCCCCGAAG GCTGAGAGCT GGCGCTGCTC GTGCCCTGTG TGCCAGACGG   60
       CGGAGCTCCG CGGCCGGACC CCGCGGCCCC GCTTTGCCAG CGACTGGAGT TTGGGGGAAG   120
       AAACTCTCCT GCGCCCCAGA AGATTTCTTC CTCGGCGAAG GGACAGCGAA AGATGAGGGT   180
       GGCAGGAAGA GAAGGCGCTT TCTGTCTGCC GGGGTCGCAG CGCGAGAGGG CAGTGCCATG   240
       TTCCTCTCCA TCCTAGTGGC GCTGTGCCTG TGGCTGCTGG CCGTGCGCGGC CGTGCGCGGC   300
       GCGCCCTGCG AGGCGGTGCG CATCCCTATG TGCCGGCACA TGCCCTGGAA CATCACGCGG   360
       ATGCCCAACC ACCTGCACCA CAGCACGCAG GAGAACGCCA TCCTGGCCAT CGAGCAGTAC   420
       GAGGAGCTGG TGGACGTGAA CTGCAGCGCC GTGCTGCGCT TCTTCTTCTG TGCCATGTAC   480
       GCGCCCATTT GCACCCTGGA GTTCCTGCAC GACCCTATCA AGCCGTGCAA GTCGGTGTGC   540
       CAACGCGCGC GCGACGACTG CGAGCCCCTC ATGAAGATGT ACAACCACAG CTGGCCCGAA   600
       AGCCTGGCCT GCGACGAGCT GCCTGTCTAT GACCGTGGCG TGTGCATTTC GCCTGAAGCC   660
       ATCGTCACGG ACCTCCCGGA GGATGTTAAG TGGATAGACA TCACACCAGA CATGATGGTA   720
       CAGGAAAGGC CTCTTGATGT TGACTGTAAA CGCCTAAGCC CCGATCGGTG CAAGTGTAAA   780
       AAGGTGAAGC CAACTTTGGC AACGTATCTC AGCAAAAAGT ACAGCTATGT TATTCGTGGA   840
       AAAATAAAAG CTGTGCAGAG GAGTGGCTGC AATGAGGTCA CAACGGTGGT GGATGTAAAA   900
       GAGATCTTCA AGTCCTCATC ACCCATCCCT CGAACTCAAG TCCCGCTCAT TACAAATTCT   960
       TCTTGCCAGT GTCCACACAT CCTGCCCCAT CAAGATGTTC TCATCATGTG TTACGAGTGG   1020
       CGTTCAAGGA TGATGCTTCT TGAAAATTGC TTAGTTGAAA AATGGAGAGA TCAGCTTAGT   1080
       AAAAGATCCA TACAGTGGGA AGAGAGGCTG CAGGAACAGC GGAGAACAGT TCAGGACAAG   1140
       AAGAAAACAG CCGGGCGCAC CAGTCGTAGT AATCCCCCCA AACCAAAGGG AAAGCCTCCT   1200
       GCTCCCAAAC CAGCCAGTCC CAAGAAGAAC ATTAAAACTA GGAGTGCCCA GAAGAGAACA   1260
       AACCCGAAAA GAGTGTGAGC TAACTAGTTT CCAAAGCGGA GACTTCCGAC TTCCTTACAG   1320
       GATGAAGGCTG GGCATTGCCT GGGACAGCCT ATGTAAGGCC ATGTGCCCCT TGCCCTAACA   1380
       ACTCACTGCA GTGCTCTTCA TAGACACATC TTGCAGCATT TTTCTTAAGG CTATGCTTCA   1440
       GTTTTTCTTT GTAAGCCATC ACAAGCCATA GTGGTAGGTT TGCCCTTTGG TACAGAAGGT   1500
       GAGTTAAAGC TGGTGGAAAA GGCTTATTGC ATTGCATTCA GAGTAACCTG TGTGCATACT   1560
       CTAGAAGAGT AGGGAAAATA ATGCTTGTTA CAATTCGACC TAATATGTGC ATTGTAAAAT   1620
       AAATGCCATA TTTCAAACAA AACACGTAAT TTTTTTACAG TATGTTTTAT TACCTTTTGA   1680
       TATCTGTTGT TGCAATGTTA GTGATGTTTT AAAATGTGAT GAAAATATAA TGTTTTTAAG   1740
       AAGGAACAGT AGTGGAATGA ATGTTAAAAG ATCTTTATGT GTTTATGGTC TGCAGAAGGA   1800
       TTTTTGTGAT GAAAGGGGAT TTTTTGAAAA ATTAGAGAAG TAGCATATGG AAAATTATAA   1860
       TGTGTTTTTT TACCAATGAC TTCAGTTTCT GTTTTTAGCT AGAAACTTAA AAACAAAAAT   1920
       AATAATAAAG AAAAATAAAT AAAAAGGAGA GGCAGACAAT GTCTGGATTC CTGTTTTTTG   1980
       GTTACCTGAT TTCCATGATC ATGATGCTTC TTGTCAACAC CCTCTTAAGC AGCACCAGAA   2040
       ACAGTGAGTT TGTCTGTACC ATTAGGAGTT AGGTACTAAT TAGTTGGCTA ATGCTCAAGT   2100
       ATTTTATACC CACAAGAGAG GTATGTCACT CATCTTACTT CCCAGGACAT CCACCCTGAG   2160
       AATAAATTTGA CAAGCTTAAA AATGGCCTTC ATGTGAGTGC CAAATTTTGT TTTTCTTCAT   2220
       TTAAATATTT TCTTTGCCTA AATACATGTG AGAGGAGTTA AATATAAATG TACAGAGAGG   2280
       AAAGTTGAGT TCCACCTCTG AAATGAGAAT TACTTGACAG TTGGGATACT TTAATCAGAA   2340
       AAAAAGAACT TATTTGCAGC ATTTTATCAA CAAATTTCAT AATTGTGGAC AATTGGAGGC   2400
       ATTTATTTTA AAAAACAATT TTATTGGCCT TTTGCTAACA CAGTAAGCAT GTATTTTATA   2460
       AGGCATTCAA TAAATGCACA ACGCCCAAAG GAAATAAAAT CCTATCTAAT CCTACTCTCC   2520
       ACTACACAGA GGTAATCACT ATTAGTATTT TGGCATATTA TTCTCCAGGT GTTTGCTTAT   2580
       GCACTTATAA AATGATTTGA ACAAATAAAA CTAGGAACCT GTATACATGT GTTTCATAAC   2640
       CTGCCTCCTT TGCTTGGCCC TTTATTGAGA TAAGTTTTCC TGTCAAGAAA GCAGAAACCA   2700
       TCTCATTTCT AACAGCTGTG TTATATTCCA TAGTATGCAT TACTCAACAA ACTGTTGTGC   2760
       TATTGGATAC TTAGGTGGTT TCTTCACTGA CAATACTGAA TAAACATCTC ACCGGAATTC
```

Seq ID NO: 64 Protein sequence
Protein Accession #: NP_003005.1

```
       1          11         21         31         41         51
       |          |          |          |          |          |
```

```
       |          |          |          |          |          |
     MFLSILVALC LWLHLALGVR GAPCEAVRIP MCRHMPWNIT RMPNHLHHST QENAILAIEQ    60
     YEELVDVNCS AVLRFFFCAM YAPICTLEFL HDPIKPCKSV CQRARDDCEP LMKMYNHSWP   120
     ESLACDELPV YDRGVCISPE AIVTDLPEDV KWIDITPDMM VQERPLDVDC KRLSPDRCKC   180
     KKVKPTLATY LSKNYSYVIH AKIKAVQRSG CNEVTIVVDV KEIFKSSSPI PRTQVPLITN   240
     SSCQCPHILP HQDVLIMCYS WRSRMMLLEN CLVEKWRDQL SKRSIQWEER LQEQRRTVQD   300
     KKKTAGRTSR SNPPKPKGKP PAPKPASPKK NIKTRSAQKR TNPKRV
```

5

Seq ID NO: 65 DNA sequence
10   Nucleic Acid Accession #: BC010423
Coding sequence: 248..1780

```
       1          11         21         31         41         51
       |          |          |          |          |          |
15   CACAGCGTGG GAAGCAGCTC TGGGGGAGCT CGGAGCTCCC GATCACGGCT TCTTGGGGGT    60
     AGCTACGGCT GGGTGTGTAG AACGGGGCCG GGGCTGGGGC TGGGTCCCCT AGTGGAGACC   120
     CAAGTGCGAG AGGCAAGAAC TCTGCAGCTT CCTGCCTTCT GGGTCAGTTC CTTATTCAAG   180
     TCTGCAGCCG GCTCCCAGGG AGATCTCGGT GGAACTTCAG AAACGCTGGG CAGTCTGCCT   240
     TTCAACCATG CCCCTGTCCC TGGGAGCCGA GATGTGGGGG CCTGAGGCCT GGCTGCTGCT   300
     GCTGCTACTG CTGGCATCAT TTACAGGCCG GTGCCCCGCG GGTGAGCTGG AGACCTCAGA   360
     CGTGGTAACT GTGGTGCTGG GCCAGGACGC AAAACTGCCC TGCTTCTACC GAGGGGACTC   420
     CGGCGAGCAA GTGGGGCAAG TGGCATGGGC TCGGGTGGAC GCGGGCGAAG GCGCCCAGGA   480
     ACTAGCGCTA CTGCACTCCA AATACGGGCT TCATGTGAGC CGGCTTACG AGGGCCGCGT    540
     GGAGCAGCCG CCGCCCCCAC GCAACCCCCT GGACGGCGTG GTGCTCCTGC GCAACGCAGT   600
     GCAGGCGGAT GAGGGCGAGT ACGAGTGCCG GGTCAGCACC TTCCCCGCCG GCAGCTTCCA   660
     GGCCCGGCTG CGGCTCCGAG TGCTGGTGCC TCCCCTGCCC TCACTGAATC CTGGTCCAGC   720
     ACTAGAAGAG GGCCAGGGCC TGACCCTGGC AGCCTCCTGC ACAGCTGAGG GCAGCCCAGC   780
     CCCCAGCGTG ACCTGGGACA CGGAGGTCAA AGGCACAACG TCCAGCCGTT CCTTCAAGCA   840
     CTCCCGGCTCT GCTGCCGTCA CCTCAGAGTT CCACTTGGTG CCTAGCCGCA GCATGAATGG   900
     GCAGCCACTG ACTTGTGTGG TGTCCCATCC TGGCCTGCTC CAGGACCAAA GGATCACCCA   960
     CATCCTCCAC GTGTCCTTCC TTGCTGAGGC CTCTGTGAGG GGCCTTGAAG ACCAAAATCT  1020
     GTGGCACATT GGCAGAGAAG GAGCCTATGCT CAAGTGCCTG AGTGAAGGGC AGCCCCCTCC  1080
     CTCATACAAC TGGACACGGC TGGATGGGCC TCTGCCCAGT GGGGTACGAG TGGATGGGGA  1140
     CACTTTGGGC TTTCCCCCAC TGACCACTGA GCACAGCGGC ATCTACGTCT GCCATGTCAG  1200
     CAATGAGTTC CCTCAAGGG ATTCTCAGGT CACTGTGGAT GTTCTTGACC CCCAGGAAGA  1260
     CTCTGGGAAG CAGGTGGACC TAGTGTCAGC CTCGGTGGTG GTGGTGGGTG TGATCGCCGC  1320
     ACTCTTGTTC TGCCTTCTGG TGGTGGTGGT GGTGCTCATG TCCCGATACC ATCGGCGCAA  1380
     GGCCCAGCAG ATGACCCAGA AATATGAGGA GGAGCTGACC CTGACCAGGG AGAACTCCAT  1440
     CCGGAGGCTG CATTCCCATC ACACGGACCC CAGGAGCCAG CCGGAGGAGA GTGTAGGGCT  1500
     GAGAGCCAGCCG GGCCACCCTG ATAGTCTCAA GGACAACAGT AGCTGCTCTG TGATGAGTGA  1560
     AGAGCCCGAG GGCCGCAGTT ACTCCACGCT GACCACGGTG AGGGAGATAG AAACACAGAC  1620
     TGAACTGCTG TCTCCAGGCT CTGGGCGGGC CGAGGAGGAG GAAGATCAGG ATGAAGGCAT  1680
     CAAACAGGCC ATGAACCATT TTGTTCAGGA GAATGGGACC CTACGGGCCA AGCCCACGGG  1740
     CAATGGCATC TACATCAATG GGCGGGGACA CCTGGTCTGA CCCAGGCCTG CCTCCCTTCC  1800
     CTAGGCCTGG CTCCTTCTGT TGACATGGGA GATTTTAGCT CATCTTGGGG GCCTCCTTAA  1860
     ACACCCCCAT TTCTTGCGGA AGATGCTCCC CATCCCACTG ACTGCTTGAC CTTTACCTCC  1920
     AACCCTTCTG TTCATCGGGA GGGCTCCACC AATTGAGTCT CTCCCACCAT GCATGCAAGT  1980
     CACTGTGTGT GTGCATGTGT GCCTGTGTGA GTGTTGACTG ACTGTGTGTG TGTGGAGGGG  2040
     TGACTGTCCG TGGAGGGGTG ACTGTGTCCG TGGTGTGTAT TATGCTGTCA TATCAGAGTC  2100
     AAGTGAACTG TGGTGTATGT GCCACGGGAT TTGAGTGGTT GCGTGGGCAA CACTGTCAGG  2160
     GTTGGCGGTG TGTGTCATGT GGCTGTGTGT GACCTCTGCC TGAAAAAGCA GGTATTTTCT  2220
     CAGACCCCAG AGCAGTATTA ATGATGCAGA GGTTGGAGGA GAGAGGTGGA GACTGTGGCT  2280
     CAGACCCAGG TGTGCGGGCA TAGCTGGAGC TGGAATCTGC CTCCGGTGTG AGGGAACCTG  2340
     TCTCCTTACCA CTTCGGAGCC ATGGGGGCAA GTGTGAAGCA GCCAGTCCCT GGGTCAGCCA  2400
     GAGGCTTGAA CTGTTACAGA AGCCCTCTGC CCTCTGGTGG CCTCTGGGCC TGCTGCATGT  2460
     ACATATTTC TGTAAATATA CATGCGCCGG GAGCTTCTTG CAGGAATACT GCTCCGAATC  2520
     ACTTTTAATT TTTTTCTTTT TTTTTCTTG CCCTTTCCAT TAGTTGTATT TTTTATTTAT  2580
     TTTTATTTTT ATTTTTTTTT AGAGTTTGAG TCCAGCCTGG ACGATATAGC CAGACCCTGT  2640
     CTGTAAAAAA ACCAAAACCC AAAAAAAAAA AAAAAAAAA
```

Seq ID NO: 66 Protein sequence
Protein Accession #: AAH10423

```
       1          11         21         31         41         51
       |          |          |          |          |          |
     MPLSLGAEMW GPEAWLLLLL LLASFTGRCP AGELETSDVV TVVLGQDAKL PCFYRGDSGR    60
     QVGQVAWARV DAGEGAQELA LLHSKYGLHV SPAYEGRVEQ PPPPRNPLDG SVLLRNAVQA   120
     DEGEYECRVS TFPAGSFQAR LRLRVLVPPL PSLNPGPALE EGQGLTLAAS CTAEGSPAPS   180
     VTWDTEVKGT TSSRSFKHSR SAAVTSEFHL VPSRSMNGQP LTCVVSHPGL LQDQRITHIL   240
     HVSFLAEASV RGLEDQNLWH IGREGAMLKC LSEGQPPPSY NWTRLDGPLP SGVRVDGDTL   300
     GFPPLTTEHS GIYVCHVSNE FSSRDEQVTV DVLDPQEDSG KQVDLVSASV VVVGVIAALL   360
     FCLLVVVVVL MSRYHRRKAQ QMTQKYEEEL TLTRENSIRR LHSHHTDPRS QPEESVGLRA   420
     EGHPDSLKDN SSCSVMSEEP EGRSYSTLTT VREIETQTEL LSPGSGRAEE EEDQDEGIKQ   480
     AMNHFVQENG TLRAKPTGNG IYINGRGHLV
```

Seq ID NO: 67 DNA sequence
Nucleic Acid Accession #: NM_001203
Coding sequence: 274..1782

```
       1          11         21         31         41         51
       |          |          |          |          |          |
     CGCGGGGCGC GGAGTCGGCG GGGCCTCGCG GGACGCGGGC AGTGCGGAGA CCGCGGCGCT    60
     GAGGACGCGG GAGCCGGGAG CGCACGCGCG GGGTGGAGTT CAGCCTACTC TTTCTTAGAT   120
     GTGAAAGGAA AGGAAGATCA TTTCATGCCT TGTTGATAAA GGTTCAGACT TCTGCTGATT   180
```

```
CATAACCATT  TGGCTCTGAG  CTATGACAAG  AGAGGAAACA  AAAAGTTAAA  CTTACAAGCC   240
TGCCATAAGT  GAGAAGCAAA  CTTCCTTGAT  AACATGCTTT  TGCGAAGTGC  AGGAAAATTA   300
AATGTGGGCA  CCAAGAAAGA  GGATGGTGAG  AGTACAGCCC  CCACCCCCCG  TCCAAAGGTC   360
TTGCGTTGTA  AATGCCACCA  CCATTGTCCA  GAAGACTCAG  TCAACAATAT  TTGCAGCACA   420
GACGGATATT  GTTTCACGAT  GATAGAAGAG  GATGACTCTG  GGTTGCCTGT  GGTCACTTCT   480
GGTTGCCTAG  GACTAGAAGG  CTCAGATTTT  CAGTGTCGGG  ACACTCCCAT  TCCTCATCAA   540
AGAAGATCAA  TTGAATGCTG  CACAGAAAGG  AACGAATGTA  ATAAAGACCT  ACACCCTACA   600
CTGCCTCCAT  TGAAAAACAG  AGATTTTGTT  GATGGACCTA  TACACCACAG  GGCTTTACTT   660
ATATCTGTGA  CTGTCTGTAG  TTTGCTCTTG  GTCCTTATCA  TATTATTTTG  TTACTTCCGG   720
TATAAAAGAC  AAGAAACCAG  ACCTCGATAC  AGCATTGGGT  TAGAACAGGA  TGAAACTTAC   780
ATTCCTCCTG  GAGAATCCCT  GAGAGACTTA  ATTGAGCAGT  CTCAGAGCTC  AGGCAAGTGGA   840
TCAGGCCTCC  CTCTGCTGGT  CCAAAGGACT  ATAGCTAAGC  AGATTCAGAT  GGTGAAACAG   900
ATTGGAAAAG  GTCGCTATGG  GGAAGTTTGG  ATGGGAAAGT  GGCGTGGCGA  AAAGGTAGCT   960
GTGAAAGTGT  TCTTCACCAC  AGAGGAAGCC  AGCTGGTTCA  GAGAGACAGA  AATATATCAG  1020
ACAGTGTTGA  TGAGGCATGA  AAACATTTTG  GGTTTCATTG  CTGCAGATAT  CAAAGGGACA  1080
GGGTCCTGGA  CCCAGTTGTA  CCTAATCACA  GACTATCATG  AAAATGGTTC  CCTTTATGAT  1140
TATCTGAAGT  CCACCACCCT  AGACGCTAAA  TCAATGCTGA  AGTTAGCCTA  CTCTTCTGTC  1200
AGTGGCTTAT  GTCATTTACA  CACAGAAATC  TTTAGTACTC  AAGGCAAACC  AGCAATTGCC  1260
CATCGAGATC  TGAAAAGTAA  AAACATTCTG  GTGAAGAAAA  ATGGAACTTG  CTGTATTGCT  1320
GACCTGGGCC  TGGCTGTTAA  ATTTATTAGT  GATACAAATG  AAGTTGACAT  ACCACCTAAC  1380
ACTCGAGTTG  GCACCAAACG  CTATATGCCT  CCAGAAGTGT  TGGACGAGAG  CTTGAACAGA  1440
AATCACTTCC  AGTCTTACAT  CATGGCTGAC  ATGTATAGTT  TTGGCCTCAT  CCTTTGGGAG  1500
GTTGCTAGGA  GATGTGTATC  AGGAGGTATA  GTGGAAGAAT  ACCAGCTTCC  TTATCATGAC  1560
CTAGTGCCCA  GTGACCCCTC  TTATGAGGAC  ATGAGGGAGA  TTGTGTGCAT  CAAGAAGTTA  1620
CGCCCCTCAT  TCCCAAACCG  GTGGAGCAGT  GATGAGTGTC  TAAGGCAGAT  GGTGAAACTC  1680
ATGACAGAAT  GCTGGGCTCA  CAATCCTGCA  TCAAGGCTGA  CAGCCCTGCG  GGTTAAGAAA  1740
ACACTTGCCA  AAATGTCAGA  GTCCCAGGAC  ATTAAACTCT  GATAGGGAGA  GAAAAGTAAG  1800
CATCTCTGCA  GAAAGCCAAC  AGGTACTCTT  CTGTTTGTGG  GCAGAGCAAA  AGACATCAAA  1860
TAAGCATCCA  CAGTACAAGC  CTTGAACATC  GTCCTGCTTC  CCAGTGGGTT  CAGACCTCAC  1920
CTTTCAGGGA  GCGACCTGGG  CAAAGACAGA  GAAGCTCCCA  GAAGGAGAGA  TTGATCCGTG  1980
TCTGTTTGTA  GGCGGAGAAA  CCGTTGGGTA  ACTTGTTCAA  GATATGATGC  AT
```

Seq ID NO: 68 Protein sequence
Protein Accession #: NP_001194

```
1           11          21          31          41          51
|           |           |           |           |           |
MLLRSAGKLN  VGTKKEDGES  TAPTPRPKVL  RCKCHHHCPE  DSVNNICSTD  GYCFTMIEED   60
DSGLPVVTSG  CLGLEGSDFQ  CRDTPIPHQR  RSIECCTERN  ECNKDLHPTL  PPLKNRDFVD  120
GPIHHRALLI  SVTVCSLLLV  LIILFCYFRY  KRQETRPRYS  IGLEQDETYI  PPGESLRDLI  180
EQSQSGSGSS  GLPLLVQRTI  AKQIQMVKQI  GKGRYGEVWM  GKWRGEKVAV  KVFFTTEEAS  240
WFRETEIYQT  VLMRHENILG  FIAADIKGTG  SWTQLYLITD  YHENGSLYDY  LKSTTLDAKS  300
MLKLAYSSVS  GLCHLHTEIF  STQGKPAIAH  RDLKSKNILV  KKNGTCCIAD  LGLAVKFISD  360
TNEVDIPPNT  RVGTKRYMPP  EVLDESLNRN  HFQSYIMADM  YSFGLILWEV  ARRCVSGGIV  420
EEYQLPYHDL  VPSDPSYEDM  REIVCIKKLR  PSFPNRWSSD  ECLRQMGKLM  TECWAHNPAS  480
RLTALRVKKT  LAKMSESQDI  KL
```

Seq ID NO: 69 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 166..1737

```
1           11          21          31          41          51
|           |           |           |           |           |
TTGGGGGTTT  ATTCTCTTCC  CTTCTAACTT  GACAGGGTCT  TGCTCTGTCA  TTCAGGCAAG   60
AGTGCAGTAG  TGTGATCACT  TCTTACTGCC  GCCTCAAGCT  TCCAGCCTCA  ACTCAAGCAA  120
TCCTCCCACC  TCAGCCACCC  AAGTGGCTGG  GACTACAGAT  TAAGAATGAC  CCAAAATAAA  180
TTAAAGCTTT  GTTCCAAAGC  CAATGTGTAT  ACTGAAGTGC  CTGATGGAGG  ATGGGGCTGG  240
GCGGTAGCTG  TTTCATTTTT  CTTCGTTGAA  GTCTTCACCT  ACGGCATCAT  CAAGACATTT  300
GGTGTCTTCT  TTAATGACTT  AATGGACAGT  TTTAATGAAT  CCAATAGCAG  GATCTCATGG  360
ATAATCTCAA  TCTGTGTGTT  TGTCTTAACA  TTTTCAGCTC  CCCTCGCCAC  AGTCCTGAGC  420
AATCGTTTCG  GACACCGTCT  GGTAGTGATG  TTGGGGGGGC  TACTTGTCAG  CACCGGGATG  480
GTGGCCGCCT  CCTTCTCACA  AGAGGTTTCT  CATATGTACG  TGGCCATCGG  CATCATCTCT  540
GGTCTGGGAT  ACTGCTTTAG  TTTTCTCCCA  ACTGTAACCA  TCCTATCACA  ATATTTTGGC  600
AAAAGACGTT  CCATAGTCAC  TGCCAGTTGCT  TCCACCAGGAG  AATGTTTCGC  TGTGTTTGCT  660
TTCGCACCAG  CAATCATGGC  TCTGAAGGAG  CGCATTGGCT  GGAGATACAG  CCTCCTCTTC  720
GTGGGCTAC   TACAGTTAAA  CATTGTCATC  TTCGGAGCAC  TGCTCAGACC  CATCTTTATC  780
AGAGGACCAG  CGTCACCGAA  AATAGTCATC  CAGGAAAATC  GGAAAGAAGC  GCAGTATATG  840
CTTGAAAATG  AGAAAACACG  AACCTCAATA  GACTCCATTG  ACTCAGGAGT  AGAACTAACT  900
ACCTCACCTA  AAAATGTGCC  TACTCACACT  AACCTGGGAC  TGGAGCCGAA  GGCCGACATG  960
CAGCAGGTCC  TGGTGAAGAC  CAGCCCCAGG  CCAAGCGAAA  AGAAAGCCCC  GCTATTAGAC  1020
TTCTCCATTT  TGAAAGAGAA  AAGTTTTATT  TGTTATGCAT  TATTTGGTCT  CTTTGCAACA  1080
CTGGGATTCT  TTGCACCTTC  CTTGTACATC  ATTCCTCTGG  GCATTAGTCT  GGGAAACCA   1140
CAGGACCGCG  CTGCTTTTTT  ATTATCTACG  ATGGCCATTG  CAGAAGTTTT  CGGAAGGATC  1200
GGAGCTGGTT  TTGTCCTCAA  CAGGGAGCCC  ATTCGTAAGA  TTTACATTGA  GCTCATCTGC  1260
GTCATCTTAT  TGACTGTGTC  TCTGTTTGCC  TTTACTTTTG  CTACTGAATT  CTGGGGTCTA  1320
ATGTCATGCA  GCATATTTTT  TGGGTTTATG  GTTGGAACAA  TAGGAGGGAC  TCACATTCCA  1380
CTGCTTGCTG  AGGATGATGT  CGTGGGCATT  GAGAAGATGT  CTTCTGCAGC  TGGGGTCTAC  1440
ATCTTCATTG  AGAGCATAGC  AGGACTGGCT  GGACCGCCCC  TTGCAGGTTT  GTTGGTGGAC  1500
CAAGTAAGA   TCTACAGCAG  GGCCTTCTAC  TCCTGCGCAG  CTGGCATGGC  CCTGGCTGCT  1560
GTGTGCCTCG  CCCTGGTGAG  ACCGTGTAAG  ATGGGACTGT  GCCAGCATCA  TCACTCAGGT  1620
GAAACAAAGG  TAGTGAGCCA  TCGTGGGAAG  ACTTACAGG   ACATACCTGA  AGACTTTCTG  1680
GAAATGGATC  TTGCAAAAA   TGAGCACAGA  GTTCACGTGC  AAATGGAGCC  GGTATGACAC  1740
ACTTTCTTAC  AACAACAGCC  ACTGTGTTGG  CTGGAGAGGG  ATGGGGTGGG  CCCAACGGGG  1800
ACACAAGGAG  GCAGAGGAGC  TAACCCCTCT  ACTCCACTTT  CAAAACTACA  TTTTAAAGGG  1860
```

```
   AATGTGTATG TGAAGAGCAC TACCAACATC GCTTTTGTTT TGTTTTGTTT TGTTTTAAGC 1920
   TTTTTTTTTT TGCTTGTTTT TAAAGCCAAA ACAAAAAACA ACCAAGCACT CTTCCATATA 1980
   TAAATCTGGC TGTATTCAGT AGCAATACAA GAGATATGTA GAAAGACTCT TTGGTTCACA 2040
 5 TTCCGATATT AAAATAGTGA CATGAACTGG CAAAGTGGTT TTAAAAGCTT TCACGTGGGA 2100
   TAAATGATTT TCTTTTTTTC TTTTCTTTCT TCCTATGGTC TTGTCTGAAT AAACTACTCT 2160
   CCTGAATAAA ACAACATCCA ACCCAGGTCA TTGAAATGAA ATTGGCCAGT C
```

Seq ID NO: 70 Protein sequence
Protein Accession #: Eos sequence

```
   1          11         21         31         41         51
   |          |          |          |          |          |
10 MTQNKLKLCS KANVYTEVPD GGWGWAVAVS PFFVEVFTYG IIKTFGVFFN DLMDSFNESN   60
   SRISWIISIC VFVLTPSAPL ATVLSNRFGH RLVVMLGGLL VSTGMVAASF SQEVSHMYVA  120
   IGIISGLGYC FSFLPTVTIL SQYFGKRRSI VTAVASTGEC FAVFAPAPAI MALKERIGWR  180
   YSLLFVGLLQ LNIVIFGALL RPIFIRGPAS PKIVIQENRK EAQYMLENEK TRTSIDSIDS  240
15 GVELTTSPKN VPTHTNLELE PKADMQQVLV KTSPRPSEKK APLLDFSILK EKSFICYALF  300
   GLFATLGPFA PSLYIIPLGI SLGIDQDRAA FLLSTMAIAE VFGRIGAGFV LNREPIRKIY  360
   IELICVILLT VSLFAPTFAT EPWGLMSCSI FFGFMVGTIG GTHIPLLAED DVVGIEKMSS  420
   AAGVYIPIQS IAGLAGPPLA GLLVDQSKIY SRAFYSCAAG MALAAVCLAL VRPCKMGLCQ  480
20 HHHSGETKVV SHRGKTLQDI PEDPLEMDLA KNEHRVHVQM EPV
```

Seq ID NO: 71 DNA sequence
Nucleic Acid Accession #: NM_004694
Coding sequence: 166..1737

```
   1          11         21         31         41         51
   |          |          |          |          |          |
   TTGGGGGTTT ATTCTCTTCC CTTCTAACTT GACAGGGTCT TGCTCTGTCA TTCAGGCAAG   60
30 AGTGCAGTAG TGTGATCACT TCTTACTGCC GCCTCAAGCT TCCAGCCTCA ACTCAAGCAA  120
   TCCTCCCACC TCAGCCACCC AAGTGGCTGG GACTACAGAT TAAGAATGAC CCAAAATAAA  180
   TTAAAGCTTC GTTCCAAAGC CAATGTGTAT ACTGAAGTGC CTGATGGAGG ATGGGGCTGG  240
   GCGGTAGCTG TTTCATTTTT CTTCGTTGAA GTCTTCACCT ACGGCATCAT CAAGACATTT  300
   GGTGTCTTCT TTAATGACTT AATGGACAGT TTTAATGAAT CCAATAGCAG GATCTCATGG  360
35 ATAATCTCAA TCTGTGTGTT TGTCTTAACA TTTTCAGCTC CCCTCGCCAC AGTCCTGAGC  420
   AATCGTTTCG GACACCGTCT GGTAGTGATG TTGGGGGGGC TACTTGTCAG CACCGGGATG  480
   GTGGCCGCCT CCTTCTCACA AGAGGTTTCT CATATGTACG TCGCCATCGG CATCATCTCT  540
   GGTCTGGGAT ACTGCTTTAG TTTTCTCCCA ACTGTAACCA TCCTATCACA ATATTTTGGC  600
   AAAAGACGTT CCATAGTCAC TGCAGTTGCT TCCACAGGAG AATGTTTCGC TGTGTTTGCT  660
40 TTCGCACCAG CAATCATGGC TCTGAAGGAG CGCATTGGCT GGAGATACAG CCTCCTCTTC  720
   GTGGGCCTAC TACAGTTAAA CATTGTCATC TTCGGAGCAC TGCTCAGACC CATCATTATC  780
   AGAGGACCAG CGTCACCGAA AATAGTCATC CAGGAAAATC GGAAAGAAGC GCAGTATATG  840
   CTTGAAAATG AGAAAACACG AACCTCAATA GACTCCATTG ACTCAGGAGT AGAACTAACT  900
   ACCTCACCTA AAAATGTGCC TACTCACACT AACCTGGAAC TGGAGCCGAA GGCCGACATG  960
45 CAGCAGGTCC TGGTGAAGAC CAGCCCCAGG CCAAGCGAAA AGAAAGCCCC GCTATTAGAC 1020
   TTCTCCATTT TGAAAGAGAA AAGTTTTATT TGTTATGCAT TATTTGGTCT CTTTGCAACA 1080
   CTGGGATTCT TTGCACCTTC CTTGTACATC ATTCCTCTGG GCATTAGTCT GGGCATTGAC 1140
   CAGGACCGCG CTGCTTTTTT ATTATCTACG ATGGCCATTG CAGAAGTTTT CGGAAGGATC 1200
   GGAGCTGGTT TTGTCCTCAA CAGGGAGCCC ATTCGTAAGA TTTACATTGA GCTCATCTGC 1260
50 GTCATCTTAT TGACTGTGTC TCTGTTTGCC TTTACTTTTG CTACTGAATT CTGGGGTCTA 1320
   ATGTCATGCA GCATATTTTT TGGGTTTATG GTTGGAACAA TAGGAGGACT CACATTCCAC 1380
   TGCTTGCTGA AGATGATGTC GTGGGCATTG CAGAAGATGT CTTCTGCAGC TGGGGTCTAC 1440
   ATCTTCATTC AGAGCATAGC AGGACTGGCT GGACCGCCCC TTGCAGGTTT GTTGGTGGAC 1500
   CAAAGTAAGA TCTACAGCAG GGCCTTCTAC TCCTGCGCAG CTGGCATGGC CCTGGCTGCT 1560
55 GTGTGCCTCG CCCTGGTGAG ACCGTGTAAG ATGGGACTGT GCCAGCGTCA TCACTCAGGT 1620
   GAAACAAAGG TAGTGAGCCA TCGTGGGAAG ACTTTACAGG ACATACCTGA AGACTTTCTG 1680
   GAAATGGATC TTGCAAAAAA TGAGCACAGA GTTCACGTGC AAATGGAGCC GGTATGACAC 1740
   ACTTTCTTAC AACAACAGCC ACTGTGTTGG CTGGAGAGGG ATGGGGTGGG CCCAACGGGG 1800
   ACACAAGGAG GCAGAGGAGC TAACCCCTCT ACTCCACTTT CAAAACTACA TTTTAAAGGG 1860
60 AATGTGTATG TGAAGAGCAC TACCAACATC GCTTTTGTTT TGTTTTGTTT TGTTTTAAGC 1920
   TTTTTTTTTT TGCTTGTTTT TAAAGCCAAA ACAAAAAACA ACCAAGCACT CTTCCATATA 1980
   TAAATCTGGC TGTATTCAGT AGCAATACAA GAGATATGTA GAAAGACTCT TTGGTTCACA 2040
   TTCCGATATT AAAATAGTGA CATGAACTGG CAAAGTGGTT TTAAAAGCTT TCACGTGGGA 2100
   TAAATGATTT TCTTTTTTTC TTTTCTTTCT TCCTATGGTC TTGTCTGAAT AAACTACTCT 2160
65 CCTGAATAAA ACAACATCCA ACCCAGGTCA TTGAAATGAA ATTGGCCAGT C
```

Seq ID NO: 72 Protein sequence
Protein Accession #: NP_004685

```
70 1          11         21         31         41         51
   |          |          |          |          |          |
   MTQNKLKLCS KANVYTEVPD GGWGWAVAVS FFFVEVFTYG IIKTFGVFFN DLMDSFNESN   60
   SRISWIISIC VFVLTPSAPL ATVLSNRFGH RLVVMLGGLL VSTGMVAASF SQEVSHMYVA  120
   IGIISGLGYC FSFLPTVTIL SQYFGKRSI VTAVASTGEC FAVFAPAPAI MALKERIGWR  180
75 YSLLFVGLLQ LNIVIFGALL RPIIIRGPAS PKIVIQENRK EAQYMLENEK TRTSIDSIDS  240
   GVELTTSPKN VPTHTNLELE PKADMQQVLV KTSPRPSEKK APLLDFSILK EKSFICYALF  300
   GLFATLGPFA PSLYIIPLGI SLGIDQDRAA FLLSTMAIAE VFGRIGAGFV LNREPIRKIY  360
   IELICVILLT VSLFAPTFAT EPWGLMSCSI FFGFMVGTIG GLTFHCLLKM MSWALQKMSS  420
   AAGVYIPIQS IAGLAGPPLA GLLVDQSKIY SRAFYSCAAG MALAAVCLAL VRPCKMGLCQ  480
80 RHHSGETKVV SHRGKTLQDI PEDPLEMDLA KNEHRVHVQM EPV
```

Seq ID NO: 73 DNA sequence
Nucleic Acid Accession #: NM_002184.1
Coding sequence: 256..3012

977

```
         1          11         21         31         41         51
         |          |          |          |          |          |
     GAGCAGCCAA AAGGCCCGCG GAGTCGCGCT GGGCCGCCCC GGCGCAGCTG AACCGGGGGC    60
     CGCGCCTGCC AGGCCGACGG GTCTGGCCCA GCCTGGCGCC AAGGGGTTCG TGCGCTGTGG   120
     AGACGCGGAG GGTCGAGGCG GCGCGGCCTG AGTGAAACCC AATGGAAAAA GCATGACATT   180
     TAGAAGTAGA AGACTTAGCT TCAAATCCCT ACTCCTTCAC TTACTAATTT TGTGATTTGG   240
     AAATATCCGC GCAAGATGTT GACGTTGCAG ACTTGGGTAG TGCAAGCCTT GTTTATTTTC   300
     CTCACCACTG AATCTACAGG TGAACTTCTA GATCCATGTG GTTATATCAG TCCTGAATCT   360
     CCAGTTGTAC AACTTCATTC TAATTTCACT GCAGTTTGTG TGCTAAAGGA AAAATGTATG   420
     GATTATTTTC ATGTAAATGC TAATTACATT GTCTGGAAAA CAAACCATTT TACTATTCCT   480
     AAGGAGCAAT ATACTATCAT AAACAGAACA GCATCCAGTG TCACCTTTAC AGATATAGCT   540
     TCATTAAATA TTCAGCTCAC TTGCAACATT CTTACATTCG GACAGCTTGA ACAGAATGTT   600
     TATGGAATCA CAATAATTTC AGGCTTGCCT CCAGAAAAAC CTAAAAATTT GAGTTGCATT   660
     GTGAACGAGG GGAAGAAAAT GAGGTGTGAG TGGGATGGTG GAAGGGAAAC ACACTTGGAG   720
     ACAAACTTCA CTTTAAAATC TGAATGGGCA ACACACAAGT TTGCTGATTG CAAAGCAAAA   780
     CGTGACACCC CCACCTCATG CACTGTTGAT TATTCTACTG TGTATTTTGT CAACATTGAA   840
     GTCTGGGTAG AAGCAGAGAA TGCCCTTGGG AAGGTTACAT CAGATCATAT CAATTTTGAT   900
     CCTGTATATA AAGTGAAGCC CAATCCGCCA CATAATTTAT CAGTGATCAA CTCAGAGGAA   960
     CTGTCTAGTA TCTTAAAATT GACATGGACC AACCCAAGTA TTAAGAGTGT TATAATACTA  1020
     AAATATAACA TTCAATATAG GACCAAAGAT GCCTCAACTT GGAGCCAGAT TCCTCCTGAA  1080
     GACACAGCAT CCACCCGATC TTCATTCACT GTCCAAGACC TTAAACCTTT TACAGAATAT  1140
     GTGTTTAGGA TTCGCTGTAT GAAGGAAGAT GGTAAGGGAT ACTGGAGTGA CTGGAGTGAA  1200
     GAAGCAAGTG GGATCACCTA TGAAGATAGA CCATCTAAAG CACCAAGTTT CTGGTATAAA  1260
     ATAGATCCAT CCCATACTCA AGGCTACAGA ACTGTACAAC TCGTGTGGAA GACATTGCCT  1320
     CCTTTTGAAG CCAATGGAAA AATCTTGGAT TATGAAGTGA CTCTCACAAG ATGGAAATCA  1380
     CATTTACAAA ATTACAACAGT TAATGCCACA AAACTGACAG TAAATCTCAC AAATGATCGC  1440
     TATCTAGCAA CCCTAACAGT AAGAAATCTT GTTGGCAAAT GACATGCAGC TGTTTTAACT  1500
     ATCCCTGCCT GTGACTTTCA AGCTACTCAC CCTGTAATGG ATCTTAAAGC ATTCCCCAAA  1560
     GATAACATGC TTTGGGTGGA ATGGACTACT CCAAGGGAAT CTGTAAAGAA ATATATACTT  1620
     GAGTGGTGTG TGTTATCAGA TAAAGCACCC TGTATCACAG ACTGGCAACA AGAAGATGGT  1680
     ACCGTGCATC GCACCTATTT AAGAGGGAAC TTAGCAGAGA GCAAATGCTA TTTGATAACA  1740
     GTTACTCCAG TATATGCTGA TGGACCAGGA AGCCCTGAAT CCATAAAGGC ATACCTTAAA  1800
     CAAGCTCCAC CTTCCAAAGG ACCTACTGTT CGGACAAAAA AAGTAGGGAA AAACGAAGCT  1860
     GTCTTAGAGT GGGACCAACT TCCTGTTGAT GTTCAGAATG GATTTATCAG AAATTATACT  1920
     ATATTTTATA GAACCATCAT TGGAAATGAA ACTGCTGTGA ATGTGGATTC TTCCCACACA  1980
     GAATATACAT TGTCCTCTTT GACTAGTGAC ACATTGTACA TGGTACGAAT GGCAGCATAC  2040
     ACAGATGAAG GTGGGAAGGA TGGTCCAGAA TTCACTTTTA CTACCCCAAA GTTTGCTCAA  2100
     GGAGAAATTG AAGCCATAGT CGTTGCCTGTT TGCTTAGCAT TCCTATTGAC AACTCTTCTG  2160
     GGAGTGCTGT TCTGCTTTAA TAAGCGAGAC CTAATTAAAA AACACATCTG GCCTAATGTT  2220
     CCAGATCCTT CAAAGAGTCA TATTGCCCAG TGGTCACCTC ACACTCCTCC AAGGCACAAT  2280
     TTTAATTCAA AAGATCAAAT GTATTCAGAT GGCAATTTCA CTGATGTAAG TGTTGTGGAA  2340
     ATAGAAGCAA ATGACAAAAA GCCTTTTCCA GAAGATCTGA AATCATTGGA CCTGTTCAAA  2400
     AAGGAAAAAA TTAATACTGA AGGACACAGC AGTGGTATTG GGGGGTCTTC ATGCATGTCA  2460
     TCTTCTAGGC CAAGCATTTC TAGCAGTGAT GAAAATGAAT CTTCACAAAA CACTTCGAGC  2520
     ACTGTCCAGT ATTCTACCGT GGTACACAGT GGCTACAGAC ACCAAGTTCC GTCAGTCCAA  2580
     GTCTTCTCAA GATCCGAGTC TACCCAGCCC TTGTTAGATT CAGAGGAGCG GCCAGAAGAT  2640
     CTACAATTAG TAGATCATGT AGATGGCGGT GATGGTATTT TGCCCAGGCA ACAGTACTTC  2700
     AAACAGAACT GCAGTCAGCA TGAATCCAGT CCAGATATTT CACATTTTGA AAGGTCAAAG  2760
     CAAGTTTCAT CAGTCAATGA GGAAGATTTT GTTAGACTTA AACAGCAGAT TTCAGACCAT  2820
     ATTTCACAAT CCTGTGGATC TGGGCAAATG AAAATGTTTC AGGAAGTTTC TGCAGCAGAT  2880
     GCTTTTGGTC CAGGTACTGA GGGACAAGTA GAAAGATTTG AAACAGTTGG CATGGAGGCT  2940
     GCGACTGATG AAGGCATGCC TAAAAGTTAC TTACCACAGA CTGTACGGCA AGGCGGCTAC  3000
     ATGCCTCAGT GAAGGACTAG TAGTTCCTGC TACAACTTCA GCAGTACCTA TAAAGTAAAG  3060
     CTAAAATGAT TTTATCTGTG AATTC
```

Seq ID NO: 74 Protein sequence
Protein Accession #: NP_002175.1

```
         1          11         21         31         41         51
         |          |          |          |          |          |
     MLTLQTWVVQ ALFIFLTTES TGKLLDPCGY ISPESFVVQL HSNFTAVCVL KEKCMDYFHV    60
     NANYIVWKTN HFTIPKEQYT IINRTASSVT FTDIASLNIQ LTCNILTFGQ LEQNVYGITI   120
     ISGLPPEKPK NLSCIVNEGK KMRCEWDGGR ETHLETNFTL KSEWATHKFA DCKAKRDTPT   180
     SCTVDYSTVY FVNIEVWVEA ENALGKVTSD HINFDPVYKV KPNPPHNLSV INSEELSSIL   240
     KLTWTNPSIK SVIILKYNIQ YRTKDASTWS QIPPEDTAST RSSFTVQDLK PFTEYVFRIR   300
     CMKEDGKGYW SDWSEASGI TYEDRPSKAP SFWYKIDPSH TQGYRTVQLV WKTLPPFEAN   360
     GKILDYEVTL TRWKSHLQNY TVNATKLTVN LTNDRYLATL TVRNLVGKSD AAVLTIPACD   420
     FQATHPVMDL KAFPKDNMLW VEWTTPRESV KKYILENCVL SDKAPCITDW QQEDGTVHRT   480
     YLRGNLAESK CYLITVTPVY ADGPGSPESI KAYLKQAPPS KGPTVRTKKV GKNEAVLEWD   540
     QLPVDVQNGF IRNYTIFYRT IIGNETAVNV DSSETEYTLS SLTSDTLYMV RMAAYTDEGG   600
     KDGPEFTFTT PKFAQGEIEA IVVPVCLAFL LTTLLGVLFC FNKRDLIKKH IWPNVPDPSK   660
     SHIAQWSPHT PPRHNFNSKD QMYSDGNFTD VSVVEIEAND KKPFPEDLKS LDLFKKEKIN   720
     TEGHSSGIGG SSCMSSSRPS ISSSDENESS QNTSSTVQYS TVVHSGYRHQ VPSVQVPSRS   780
     ESTQPLLDSE ERPEDLQLVD HVDGGDGILP RQQYFKQNCS QHESSPDISH FERSKQVSSV   840
     NEEDFVRLKQ QISDHISQSC GSGQMKMFQE VSAADAFGPG TEGQVERFET VGMEAATDEG   900
     MPKSYLPQTV RQGGYMPQ
```

Seq ID NO: 75 DNA sequence
Nucleic Acid Accession #: NM_022131
Coding sequence: 11..2878

```
         1          11         21         31         41         51
         |          |          |          |          |          |
     TGCTGCGAGG ATGCTGCCTG GGCGGCTGTG CTGGGTGCCG CTCCTGCTGG CGCTGGGCGT    60
```

```
       GGGGAGCGGC AGCGGCGGTG GCGGGGACAG CCGGCAGCGC CGCCTCCTCG CGGCTAAAGT  120
       CAATAAGCAC AAGCCATGGA TCGAGACTTC ATATCATGGA GTCATAACTG AGAACAATGA  180
       CACAGTCATT TTGGACCCAC CACTGGTAGC CCTGGATAAA GATGCACCGG TTCCTTTTGC  240
       AGGGGAAATC TGTGCGTTCA AGATCCATGG CCAGGAGCTG CCCTTTGAGG CTGTGGTGCT  300
 5     CAACAAGACA TCAGGAGAGG GCCGGCTCCG TGCCAAGAGC CCCATTGACT GTGAGTTGCA  360
       GAAGGAGTAC ACATTCATCA TCCAGGCCTA TGACTGTGGT GCTGGGCCCC ACGAGACAGC  420
       CTGGAAAAAG TCACACAAGG CCGTGGTCCA TATACAGGTG AAGGATGTCA ACGAGTTTGC  480
       TCCCACCTTC AAAGAGCCAA CCTACAAGGC TGTTGTGACG GAGGGCAAGA TCTATGACAG  540
       CATTCTGCAG GTGGAGGCCA TTGACGAGGA CTGCTCCCCA CAGTACAGCC AGATCTGCAA  600
10     CTATGAAATC GTCACCACAG ATGTGCCTTT TGCCATCGAC AGAAATGGCA ACATCAGGAA  660
       CACTGAGAAG CTGAGCTATG ACAAACAACA CCAGTATGAG ATCCTGGTGA CCGCCTACGA  720
       CTGTGGACAG AAGCCCGCTG CTCAGGACAC CCTGGTGCAG GTGGATGTGA AGCCAGTTTG  780
       CAAGCCTGGC TGGCAAGACT GGACCAAGAG GATTGAGTAC CAGCCTGGCT CCGGGGAGCAT  840
       GCCCCTGTTC CCCAGCATCC ACCTGGAGAC GTGCGATGGA GCCGTGTCTT CCCTCCAGAT  900
15     CGTCACAGAG CTGCAGACTA ATTACATTGG GAAGGGTTGT GACCGGGGAGA CCTACTCTGA  960
       GAAATCCCTT CAGAAGTTAT GTGGAGCCTC CTCTGGCATC ATTGACCTCT TGCCATCCCC  1020
       TAGCGCTGCC ACCAACTGGA CTGCAGGACT GCTGGTGGAC AGCAGTGAGA TGATCTTCAA  1080
       GTTTGACGCC AGGCAGGGTG CCAAAATCCC CGATGGGATT GTGCCCAAGA ACCTGACCGA  1140
       TCAGTTCACC ATCACCATGT GGATGAAACA CGGCCCCAGC CCTGGTGTGA GAGCCGAGAA  1200
20     GGAAACCATC CTCTGCAACT CAGACAAAAC CGAAATGAAC CGGCATCACT ATGCCCTGTA  1260
       TGTGCACAAC TGCCGCCTCG TCTTTCTCTT GCGGGAAGGAC TTCGACCAGG CTGACACCTT  1320
       TCGCCCCGCG GAGTTCCACT GGAAGCTGGA TCAGATTTGT GACAAAGAGT GGCACTACTA  1380
       TGTCATCAAT GTGGAGTTTC CTGTGGTAAC CTTATACATG GATGGAGCAA CATATGAACC  1440
       ATACCTGGTG ACCAACGACT GGCCCATTCA TCCATCTCAC ATAGCCATGC AACTCACAGT  1500
25     CGGCGCTTGT TGGCAAGGAG GAGAAGTCAC CAAACCACAG TTTGCTCAGT TCTTTCATGG  1560
       AAGCCTGGCC AGTCTCACCA TCCGCCCTGG CAAAATGGAA AGCCAGAAGG TGATCTCCTG  1620
       CCTGCAGGCC TGCAAGGAAG GGCTGGACAT TAATTCCTTG GAAAGCCTTG GCCAAGGAAT  1680
       AAAGTATCAC TTCAACCCCT CGCAGTCCAT CCTGGTGATG GAAGGTGACG ACATTGGGAA  1740
       CATTAACCGT GCTCTCCAGA AAGTCTCCTA CATCAACTCC AGGCAGTTCC CAACGGCGGG  1800
30     TGTGCGGCGC CTCAAAGTAT CCTCCAAAGT CCAGTGCTTT GGGGAAGACG TATGCATCAG  1860
       TATCCCTGAG GTAGATGCCT ATGTGATGGT CCTCCAGGCC ATCGAGCCCC GGATCACCCT  1920
       CCGGGGCACA GACCCACTTCT GGAGACCTGC TGCCCAGTTT GAAAGTGCCA GGGGAGTGAC  1980
       CCTCTTCCCT GATATCAAGA TTGTGAGCAC CTTCGCCAAA ACCGAAGCCC CCGGGGACGT  2040
       GAAAACCACA GACCCCAAAT CAGAAGTCTT AGAGGAAATG CTTCATAACT TAGATTTCTG  2100
35     TGACATTTTG GTGATCGGAG GGGACTTGGA CCCAAGGCAG GAGTGCTTGG AGCTCAACCA  2160
       CAGTGAGCTC CACCAACGAC ACCTGGATGC CACTAATTCT ACTGCAGGCT ACTCCATCTA  2220
       CGGTGTGGGC TCCATGAGCC GCTATGAGCA GGTGCTACAT CACATCCGCT ACCGCAACTG  2280
       GCGTCCGGCT TCCCTTGAGG CCCGGCGTTT CCGGATTAAG TGCTCAGAAC TCAATGGGCG  2340
       CTACACTAGC AATGAGTTCA ACTTGGAGGT CAGCATCCTT CATGAAGACC AAGTCTCAGA  2400
40     TAAGGAGCAT GTCAATCATC TGATTGTGCA GCCTCCCTTC CTCCAGTCTG TCCATCATCC  2460
       TGAGTCCCGG AGTAGCATCC AGCACAGTTC AGTGGTCCCA AGCATTGCCA CAGTGGTCAT  2520
       CATCATCTCC GTGTGCATGC TTGTGTTTGT CGTGGCCATG GGTGTGTACC GGGTCCGGAT  2580
       CGCCCACCAG CACTTCATCC AGGAGACTGA GGCTGCCAAG GAATCTGAGA TGGACTGGGA  2640
       CGATTCTGCG CTGACTATCA CAGTCAACCC CATGGAGAAA CATGAAGGAC CAGGGCATGG  2700
45     GGAAGATGAG ACTGAGGGAG AAGAGGAGGA AGAAGCCGAG GAAGAAATGA GCTCCAGCAG  2760
       TGGCTCTGAC GACAGCGAAG AGGAGGAGGA GGAGGAAGGG ATGGGCAGAG GCAGACATGG  2820
       GCAGAATGGA GCCAGGCAAG CCCAGCTGGA GTGGGATGAC TCCACCCTCC CCTACTAGTG  2880
       CCCAGGGGTC TGCTGCCTGG CCCACATGTC CCTTTTGTAA ACCCTGACCC AGTGTATGCC  2940
50     CATGTCTATC ATACCTCACC TCTGATGTCT GTGACATGTC TGGGAAGGCC TTCTCCAGCT  3000
       TCCTGGAGCC CACCCTTTAA GCCTTGGGCA CTCCCTGTGT TTCATCCATG GGGAAGTTCC  3060
       AAGAAGCCCA GCATGGCCAT CAGTGAGGAC TTCAGGGTAG ACTTTGTCCT GTAGCCTCCA  3120
       CTTCTGCCCT AAGTTCCCCA GCATCCTGAC TACCTGTCTG CAGAGTTTGC CTTTGTTTTT  3180
       TCCTGCAGGG AAGAAGGCCC ACCTTTGTGT CACTCACCTC CCCAGGCCTCA GAGTCCCCAA  3240
       GGCCCTGGGG TTCCAACTCA CTGTGCGTCT CCTCCACACA GACCAGTAGG TTCTCCTATG  3300
55     CTGACTCCAG GTTGCTTCAT ACAAGGAGGG TGGTTGAACT TCACACACGT AAGGTCTTAG  3360
       TGCTTAACAG TTTAAAGGAA AGTCCTTGTT GAGGCAGAAC TAAGTTTACA GGGAAAGGTA  3420
       CACACATTCT CTCTCTCTCT CTCTCTCTGT CTATCTAGTT CCCCAGCTTG GAGAGCCTTT  3480
       CCCCTTGCTT CTTTCTGAGG CCATATAAGC TTATAAGAAA AGTCCCAAAC CAAGAATAGG  3540
       TCCTTGGCCA CAAGCAGGGT CTGATCCCCC ATCAGAGCTA TCTGAGCCTG CCTGTCTGGA  3600
60     CACCTGCTGC AACCATGCAG CTACCCTGCC AGGGGCACTC AGCAAACAGA ACCACAGGGC  3660
       CCAGGAGGCA TTCCACACAG GCACTGCCCC AGGACAACAC AACAAGGACA GTCACAACAA  3720
       GGACAACAAG GACACAACAC AACACACAAC AAGGACAGTC ACAACAAGCC ACAACAAGCC TAGAGCCAGA  3780
       AAGCAGATGG AAATGCTAAT GAGGTCAAAC GTAGGCTTCA TGGTGGGGTG AGTGGGGGTG  3840
       GCTGGGCTCC CCCAGGACAG AGGGGACCCT GAGGTTGGCA AGGCTCTCAC CACTCAGCCT  3900
65     TATGGTCCCT TATCTCCTAT CTTCCCTCTT GAGAAAATAC ACGCTTTCTG CATGTATTAG  3960
       AAACGCACGA GCTCCACCAA GTCTACAATG AAAGTTTGAA ATTTAACTGC AAGGAATTAG  4020
       AAGCATATTT GCAATCATTG CAGCTTCTTC TTTCTTCTGC TCATAAAAGG AGGAACACTT  4080
       TAGATAGAGG GCAAATATAT CTGAAAACCT AATTTCTTTC TTTTTTTGAT AAGGAAATCT  4140
       TTTCCATCTC CATCCTAACA TGCACAACCT GTGAAGAGAA TTGTTTCTAT AGTAACTGGT  4200
70     CTGTGATCTT TTGTGGCCAA GAGAATAGCA GGCAAGAATT AGGGGCCTTGA CAGAATTTCC  4260
       ACGAAGCTCT GAGAACATGT TTGTTTCGAA TGTCTGATTC CTCTTTGTCA TCAATGTGTA  4320
       TGCTCTGTCC CCATCCTTCA CTCCTCCTCA AGCTCACACC AATTGGTTTG GCACAGGCAC  4380
       AGAGCTGGTC CCTAGTTAAG TGGCATTTAT GTTAAAAAAA A
```

75   Seq ID NO: 76 <u>Protein sequence</u>
       Protein Accession #: NP_071414

```
       1          11         21         31         41         51
       |          |          |          |          |          |
80     MLPGRLCWVP LLLALGVGSG SGGGGDSRQR RLLAAKVNKH KPWIETSYHG VITENNDTVI  60
       LDPPLVALDK DAPVPPAGEI CAFKIHGQEL PFEAVVLNKT SGEGRLRAKS PIDCELQKEY  120
       TFIIQAYDCG AGPHETAWKK SHKAVVHIQV KDVNEFAPTF KEPAYKAVVT EGKIYDSILQ  180
       VEAIDEDCSP QYSQICNYEI VTTDVPFAID RNGNIRNTEK LSYDKQHQYE ILVTAYDCGQ  240
       KPAAQDTLVQ VDVKPVCKPG WQDWTKRIEY QPGSGSMPLF PSIHLETCDG AVSSLQIVTE  300
```

```
         LQTNYIGKGC  DRETYSEKSL  QKLCGASSGI  IDLLPSPSAA  TNWTAGLLVD  SSEMIFKFDG  360
         RQGAKIEDGI  VPKNLTDQFT  ITNWMKHGPS  PGVRAEKETI  LCNSDKTEMN  RHHYALYVHN  420
         CRLVFLLRKD  FDQADTFRPA  EFHWKLDQIC  DKEWHYYVIN  VEPPVVTLYM  DGATYEPYLV  480
         TNDWPIHPSH  IAMQLTVGAC  WQGGEVTKPQ  FAQPFHGSLA  SLTIRPGKME  SQKVISCLQA  540
         CKEGLDINSL  ESLGQGIKYH  FNPSQSILVM  BGDDIGNINR  ALQKVSYINS  RQFPTAGVRR  600
         LRVSSKVQCF  GEDVCISIPE  VDAYVMVLQA  IEPRITLRGT  DHFWRPAAQF  ESARGVTLFP  660
         DIKIVSTFAK  TEAPGDVKTT  DPKSEVLEEM  LHNLDFCDIL  VIGGDLDPRQ  ECLELNESEL  720
         HQRHLDATNS  TAGYSIYGVG  SMSRYEQVLH  HIRYRNWRPA  SLEARRFRIK  CSELNGRYTS  780
         NEFNLEVSIL  HEDQVSDKEE  VNHLIVQPPP  LQSVHHPESR  SSIQHSSVVP  SIATVVIIIS  840
         VCMLVFVVAM  GVYRVRIAEQ  HFIQETEAAK  ESEMDWDDSA  LTITVNPMEK  HEGPGHGEDE  900
         TEGEEEEEAE  REMSSSSGSD  DSEEEEEEEG  MGRGREGQNG  ARQAQLEWDD  STLPY
```

Seq ID NO: 77 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 482__3007

```
         1          11         21         31         41         51
         |          |          |          |          |          |
         AACTGAGCTA  ACAAGAAATA  CTAGAAAAGG  AGGAAGGAGA  ACATTGCTGC  AGCTTGGATC  60
         TACAACCTAA  GAAAGCAAGA  GTGATCAATC  TCAGCTCTGT  TAAACATCTT  GTTTACTTAC  120
         TGCATTCAGC  AGCTTGCAAA  TGGTTAACTA  TATGCAAAAA  AGTCAGCATA  GCTGTGAAGT  180
         ATGCCGTGAA  TTTTAATTGA  GGGAAAAAGG  GACAATTGCT  TCAGGATGCT  CTAGTATGCA  240
         CTCTGCTTGA  AAATATTTTCA ATGAAATGCT  CAGTATTCTA  TCTTTGACCA  GAGGTTTTAA  300
         CTTTATGAAG  CTATGGGACT  TGACAAAAAG  TGATATTTGA  GAAGAAAGTA  CGCAGTGGTT  360
         GGTGTTTTCT  TTTTTTTAAT  AAAGGAATTG  AATTACTTTG  AACACCTCTT  CCAGCTGTGC  420
         ATTACAGATA  ACGTCAGGAA  GAGTCTCTGC  TTTACAGAAT  CGGATTTCAT  CACATGACAA  480
         CATGAAGCTG  TGGATTCATC  TCTTTTATTC  ATCTCTCCTT  GCCTGTATAT  CTTTACACTC  540
         CCAAACTCCA  GTGCTCTCAT  CCAGAGGCTC  TTGTGATTCT  CTTTGCAATT  GTGAGGAAAA  600
         AGATGGCACA  ATGCTAATAA  ATTGTGAAGC  AAAAGGTATC  AAGATGGTAT  CTGAAATAAG  660
         TGTGCCACCA  TCACGACCTT  TCCAACTAAG  CTTATTAAAT  AACGGCTTGA  CGATGCTTCA  720
         CACAAATGAC  TTTTCTGGGC  TTACCAATGC  TATTTCAATA  CACCTTGGAT  TTAACAATAT  780
         TGCAGATATT  GAGATAGGTG  CATTTAATGG  CCTTGGCCTC  CTGAAACAAC  TTCATATCAA  840
         TCACAATTCT  TTAGAAATTC  TTAAAGAGGA  TACTTTCCAT  GGACTGGAAA  ACCTGGAATT  900
         CCTGCAAGCA  GATAACAATT  TTATCACAGT  GATTGAACCA  AGTGCCTTTA  GCAAGCTCAA  960
         CAGACTCAAA  GTGTTAATTT  TAAATGACAA  TGCTATTGAG  AGTCTTCCTC  CAAACATCTT  1020
         CCGATTTGTT  CCTTTAACCC  ATCTAGATCT  TCGTGGAAAT  CAATTACAAA  CATTGCCTTA  1080
         TGTTGGTTTT  CTCGAACACA  TTGGCCGAAT  ATTGGATCTT  CAGTTGGAGG  ACAACAAATG  1140
         GGCCTGCAAT  TGTGACTTAT  TGCAGTTAAA  AACTTGGTTG  GAGAACATGC  CTCCACAGTC  1200
         TATAATTGGT  GATGTTGTCT  GCAACAGCCC  TCCATTTTTT  AAAGGAAGTA  TACTCAGTAG  1260
         ACTAAAGAAG  GAATCTATTT  GCCCTACTCC  ACCAGTGTAT  GAAGAACATG  AGGATCCTTC  1320
         AGGATCATTA  CATCTGGCAG  CAACATCTTC  AATAAATGAT  AGTCGCATGT  CAACTAAGAC  1380
         CACGTCCATT  CTAAAACTAC  CCACCAAAGC  ACCAGGTTTG  ATACCTTATA  TTACAAAGCC  1440
         ATCCACTCAA  CTTCCAGGAC  CTTACTGCCG  TATTCCTTGT  AACTGCAAAG  TCCTATCCCC  1500
         ATCAGGACTT  CTAATACATT  GTCAGGAGCG  CAACATTGAA  AGCTTATCAG  ATCTGAGACC  1560
         TCCTCCGCAA  AATCCTAGAA  AGCTCATTCT  AGCGGGAAAT  ATTATTCACA  GTTTAATGAA  1620
         GTCTGATCTA  GTGGAATATT  TCACTTTGGA  AATGCTTCCA  TTGGGAAACA  ATCGTATTGA  1680
         AGTTCTTGAA  GAAGGATCGT  TTATGAACCT  AACGAGATTA  CAAAAACTCT  ATCTAAATGG  1740
         TAACCACCTG  ACCAATTAA   GTAAAGGCAT  GTTCCTTGGT  CTCCATAATC  TTGAATACTT  1800
         ATATCTTGAA  TACAATGCCA  TTAAGGAAAT  ACTGCCAGGA  ACCTTTAATC  CAATGCCTAA  1860
         ACTTAAAGTC  CTGTATTTAA  ATAACAACCT  CCTCCAAGTT  TTACCACCAC  ATATTTTTTC  1920
         AGGGGTTCCT  CTAACTAAGG  TAAATCTTAA  AACAAACCAG  TTTACCCATC  TACCTGTAAG  1980
         TAATATTTTG  GATGATCTTG  ATTTACTAAC  CCAGATTGAC  CTTGAGGATA  ACCCCTGGGA  2040
         CTGCTCCTGT  GACCTGGTTG  GACTGCAGCA  ATGGATACAA  AAGTTAAGCA  AGAACACAGT  2100
         GACAGATGAC  ATCCTCTGCA  CTTCCCCCGG  GCATCTCGAC  AAAAAGGAAT  TGAAAGCCCT  2160
         AAATAGTGAA  ATTCTCTGTC  CAGGTTTAGT  AAATAACCCA  TCCATGCCAA  CACAGACTAG  2220
         TTACCTTATG  GTCACCACTC  CTGCAACAAC  AACAAATACG  GCTGATACTA  TTTTACGATC  2280
         TCTTACGGAC  GCTGTGCCAC  TGTCTGTTCT  AATATTGGGA  CTTCTGATTA  TGTTCATCAC  2340
         TATTGTTTTC  TGTGCTGCAG  GGATAGTGGT  TCTTGTTCTT  CACCGCAGGA  GAAGATACAA  2400
         AAAGAAACAA  GTAGATGAGC  AAATGAGAGA  CAACAGTGCT  GTGCATCTTC  AGTAGCAGCAT 2460
         GTATGGCCAT  AAAACCACTC  ATCACACTAC  TGAAAGACCC  TCTGCCTCAC  TCTATGAACA  2520
         GCACATGGTG  AGCCCCATGG  TTCATGTCTA  TAGAAGTCCA  TCCTTTGGTC  CAAAGCATCT  2580
         GGAGAGGAGA  GAAGAGAGGA  ATGAGAAAGA  AGGAAGTGAT  GCAAAACATC  TCCAAAGAAG  2640
         TCTTTTGGAA  CAGGAAAATC  ATTCACCACT  CACAGGGTCA  AATATGAAAT  ACAAAACCAC  2700
         GAACCAATCA  ACAGAATTTT  TATCCTTCCA  AGATGCCAGC  TCATTGTACA  GAAACATTTT  2760
         AGAAAAAGAA  AGGGAACTTC  AGCAACTGGG  AATCACAGAA  TACCTAAGGA  AAAACATTGC  2820
         TCAGCTCCAG  CCTGATATGG  AGGCACATTA  TCCTGGAGCC  CACCAAGAGC  TGAAGTTAAT  2880
         GGAAACATTA  ATGTACTCAC  GTCCAAGGAA  GGTATTAGTG  GAACAGACAA  AAAATGAGTA  2940
         TTTTGAACTT  AAAGCTAATT  TACATGCTGA  ACCTGACTAT  TTAGAAGTCC  TGGAGCAGCA  3000
         AACATAGATG  GAGAGTTTGA  GGGCTTTCGC  AGAAATGCTG  TGATTCTGTT  TTAAGTCCAT  3060
         ACCTTGTAAA  TAAGTGCCTT  ACGTGAGTGT  GTCATCAATC  AGAACCTAAG  CACAGCAGTA  3120
         AACTATGGGG  AAAAAAAAAG  AAGAAGAAAA  GAAACTCAGG  GATCACTGGG  AGAAGCCATG  3180
         GCATTATCTT  CAGGCAATTT  AGTCTGTCCC  AAATAAAATC  AATCCTTGCA  TGTAAATC
```

Seq ID NO: 78 Protein sequence
Protein Accession #: Eos sequence

```
         1          11         21         31         41         51
         |          |          |          |          |          |
         MKLWIHLFYS  SLLACISLHS  QTPVLSSRGS  CDSLCNCEEK  DGTMLINCEA  KGIKMVSEIS  60
         VPPSRPPQLS  LLNNGLTMLH  TNDFSGLTNA  ISIHIGFNNI  ADIEIGAFNG  LGLLKQLHIN  120
         HNSLEILKED  TFHGLENLEF  LQADNNFITV  IEPSAFSKLN  RLKVLILNDN  AIESLPPNIF  180
         RFVPLTHLDL  RGNQLQTLPY  VGFLEHIGRI  LDLQLEDNKW  ACNCDLLQLK  TWLENMPPQS  240
         IIGDVVCNSP  PFFKGSILSR  LKKESICPTP  PVYEEHEDPS  GSLHLAATSS  INDSRMSTKT  300
         TSILKLPTKA  PGLIPYITKP  STQLPGPYCP  IPCNCKVLSP  SGLLIHQQER  NIESLSDLRP  360
```

```
      PPQNPRKLIL AGNIIHSLMK SDLVEYPTLE MLHLGNNRIE VLEEGSFMNL TRLQKLYLNG    420
      NHLTKLSKGM PLGLHNLBYL YLEYNAIKEI LPGTFNPMPK LKVLYLNNNL LQVLPPHIFS    480
      GVPLTKVNLK TNQFTHLPVS NILDDLDLLT QIDLEDNPWD CSCDLVGLQQ WIQKLSKNTV    540
      TDDILCTSPG HLDKKELKAL NSEILCPGLV NNPSMPTQTS YLMVTTPATT TNTADTILRS    600
      LTDAVPLSVL ILGLLIMFIT IVFCAAGIVV LVLHRRRRYK KKQVDEQMRD NSPVHLQYSM    660
      YGHKTTHHTT ERPSASLYEQ HMVSPMVHVY RSPSFGPKHL EEEEERNEKE GSDAKHLQRS    720
      LLEQENHSPL TGSNMKYKTT NQSTEPLSFQ DASSLYRNIL EKERELQQLG ITEYLRKNIA    780
      QLQPDMEAHY PGAHEELKLM ETLMYSRPRK VLVEQTKNEY FELKANLHAE PDYLEVLEQQ    840
      T
```

      Seq ID NO: 79 DNA sequence
      Nucleic Acid Accession #: NM_016640.2
      Coding sequence: 39..1358

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      GCTTAAGTTG ACCTCTGGGT CCGGAATCGC GGGCAAAGAT GGCGGCGGCC AGGTGTTGGA    60
      GGGCCTTTGCT ACGCGGTCCG AGGCTTTCAT TGCACACCGC GGCTAATGCC GCCGCCACGG   120
      CTACAGAAAC GACCTCCCAA GACGTCGCGG CGACCCCCGT CGCGCGGTAC CCGCCGGATTG   180
      TGGCCTCCAT GACAGCCGAC AGCAAAGCTG CACGGCTGCG GCGGATCHAG CGCTGGCAGG    240
      CGACGGTGCA CGCTGCGGAG TCGGTAGACG AGAAGCTGCG AATCCTCACC AAGATGCAGT    300
      TTATGAAGTA CATGGTTTAC CCGCAGACCT TCGCGCTGAA TGCCGACCGC TGGTACCAGT    360
      ACTTCACCAA GACCGTGTTC CTGTCGGGTC TGCCGCCGCC CCCAGCGGAG CCCGAGCCCG    420
      AGCCCGAACC CGAACCTGAA CCTGCGCTGG ACCTCGCGGC GCTGCGTGCG GTCGCCTGCG    480
      ACTGCCTGCT GCAGGAGCAC TTCTACCTGC GGCGCAGGCG GCGCGTGCAC CGTTACGAGG    540
      AGAGCGAGGT CATATCTTTG CCCTTCCTGG ATCAGCTGGT GTCAACCCTC GTGGGCCTCC    600
      TCAGCCCACA CAACCCGGCC CTGGCCGCTG CGGCCCTGGTCA TTATAGATGC CCAGTTCATT    660
      TTTACTGGGT GCGTGGTGAA GAAATTATTC CTCGTGGTCA TCGAAGAGGT CGAATTGATG    720
      ACTTGCGATA CCAGATAGAT GATAAACCAA ACAACCAGAT TCGAATATCC AAGCAACTCG    780
      CAGAGTTTGT GCCATTGGAT TATTCTGTTC CTATAGAAAT CCCCACTATA AAATGTAAAC    840
      CAGACAAACT TCCATTATTC AAACGGCAGT ATGAAAACCA CATATTTGTT GGCTCAAAAA    900
      CTGCAGATCC TTGCTGTTAC GGTCACACCC AGTTTCATCT GTTACCTGAC AAATTAAGAA    960
      GGGAAAGGCT TTTGAGACAA AACTGTGCTG ATCAGATAGA AGTTGTTTTT AGAGCTAATG   1020
      CTATTGCAAG CCTTTTTGCT TGGACTGGAG CACAAGCTAT GTATCAAGGA TTCTGGAGTG   1080
      AAGCAGATGT TACTCGACCT TTTGTCTCCC AGGCTGTGAT CACAGATGGA AAATACTTTT   1140
      CCTTTTTCTG CTACCAGCTA AATACTTTGG CACTGACTAC ACAAGCTGAT CAAAATAACC   1200
      CTCGTAAAAA TATATGTTGG GGTACACAAA GTAAGCCTAT TTATGAAACA ATTGAGGATA   1260
      ATGATGTGAA AGGTTTTAAT GATGATGTTC TACTTCAGAT AGTTCACTTT CTACTGAATA   1320
      GACCAAAAGA AGAAAAATCA CAGCTGTTGG AAAACTGAAA AAGCATATTT GATTGAGAAC   1380
      TGTGGGAATA TTTAAATTTT ACTGAAGGAA CAATAATGAT GAGATTTGTA ACTGTCAACT   1440
      ATTAAATACA TTGATTTTTG AGACAAAAAA AAAAAAAAAA AA
```

      Seq ID NO: 80 Protein sequence
      Protein Accession #: NP_057724.1

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      MAAARCWRPL LRGPRLSLHT AANAAATATE TTSQDVAATP VARYPPIVAS MTADSKAARL    60
      RRIERWQATV HAAESVDEKL RILTKMQFMK YMVYPQTFAL NADRWYQYFT KTVFLSGLPP    120
      PPAEPEPEPE PEPEPALULA ALRAVACDCL LQEHFYLRRR RVHRYEESE VISLPFLDQL    180
      VSTLVGLLSP HNPALAAAAL DYRCPVHFYW VRGEEIIPRG HRRGRIDDLR YQIDDKPNNQ    240
      IRISKQLAEF VPLDYSVPIE IPTIKCKPDK LPLFKRQYEN HIFVGSKTAD PCCYGHTQFH    300
      LLPDKLRRER LLRQNCADQI EVVFRANAIA SLFAWTGAQA MYQGFWSEAD VTRPFVSQAV    360
      ITDGKYFSFF CYQLNTLALT TQADQNNPRK NICWGTQSKP LYETIEDNDV KGFNDDVLLQ    420
      IVHFLLNRPK EEKSQLLEN
```

      Seq ID NO: 81 DNA sequence
      Nucleic Acid Accession #: FGENESH predicted
      Coding sequence: 1..2070

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      ATGAGCGGTG CGGGGGTGGC GGCTGGGACG CGGCCCCCCA GCTCGCCGAC CCCGGGCTCT    60
      CGGCGCCGGC GCCAGCGCCC CTCTGTGGGC GTCCAGTCCT·TGAGGCCGCA GAGCCCGCAG   120
      CTCAGGCAGA GCGACCCGCA GAAACGGAAC CTGGACCTGG AGAAAAGCCT GCAGTTCCTG   180
      CAGCAGCAGC ACTCGGAGAT GCTGGCCAAG CTCCATGAGG AGATCGAGCA TCTGAAGCGG   240
      GAAAACAAGG GTGAGCCGGC GCGGGGCCCT AGGCCGGCCC TGCCTCCCCA GGCACACTCA   300
      ACACTGCCGC TCCCGCAGCA CAGAAACACA GCCATCAACT CCAGCACACG CCTGGGCTCA   360
      GGGGGAACAC AGGACGGGGA GCCCCTCCAG ACTGTCCTTG CCCACCTGGC TGCACTGGCC   420
      CCTGTATGCC AACCCAGTGG GTACAGGTTC TGGGGGACCT GGACAGATGC CGCTACCTCT   480
      AGCCGTGGCT GGACGATGTT ATGCAGCCAA GCACAGCACG TGCTGCTCTC GGGAAGCCCA   540
      GGGCCTGAGG TCATTGCAGG GCGGCAGGTG GCCACAGGGT GCTCCCCAGA CCTCCCTCCT   600
      CCAAGTAGAG CTGAAATGGG AAGGAACCCC TGGGACAGCC CCTGCCCTGC TAGATCTTTG   660
      CCTCAGATTG CTGCTGTGGC CAGGCCCAGG ATTTCCAGCC CTATGGCTCT GAGTCCTCAC   720
      ATGCTGGGGG CCCAGGGGAT ATGGACACAC TCCATCCAGG GATCCCTTCC TGCCATCTGG   780
      GCAGCAACCA TGGGGACAAA GGGAGGAAGC AGAGTCCTGT TTCCTTGCCA CTTGTCCAAG   840
      GCACTTCCCC ATCCTGACAG CGGCCCCCAC CCAGCCCAGG ATCCTGGGCT GTGGTCTCAA   900
      GCTCACTTCC CATTATCTTT GGGGCTGGGG CTGACATCAG GAGGACATCT GACTGGTGGA   960
      TGGAGCCAGC CTGGGAACAT CGCAGCTGGG GCAGTGCCTA GGGCTCTCCC TTCCCAGGGA  1020
      GACATGGAGA AGGGGGTTGA GGGAGGGCCC TTCCCTAGCC GCTGTGGCAA CTCCAGTGAG  1080
      CTGTTCTGGG CAAAGTGTGG CCCAAGTCGG CAGCCCCAGC CCTGCAGTGC TGGGGACCCT  1140
      GACAGGACAC GGGAAGAGGC CATGCTTTCC CTCGGGACCT GCTGTTCCAT GTGTCCCAAG  1200
      CCCTCCTGCT TTCCAGATGG CCCCTCAGGA AACCACCTTT CCAGGGCCTC TGCTCCCTTG  1260
      GGCGCTCGCT GGGTCTGCAT CAACGGAGTG TGGGTAGAGC CGGGAGGACC CAGCCCTGCC  1320
```

```
      AGGCTGAAGG AGGGCTCCTC ACGGACACAC AGGCCAGGAG GCAAGCGTGG GCGTCTTGCG   1380
      GGCGGTAGCG CCGACACTGT GCGCTCTCCT GCAGACAGCC TCTCCATGTC AAGCTTCCAG   1440
      TCTGTCAAGT CCATCTCTAA TTCAGCCAAC TCTCAAGGCA AGGCCAGGCC CCAGCCCGGC   1500
      TCCTTCAACA AGCAAGATTC AAAAGCTGAC GTCTCCCAGA AGGCGGACCT GGAAGAGGAG   1560
  5   CCCCTACTTC ACAACAGCAA GCTGGACAAA GTTCCTGGGG TACAAGGGCA GGCCAGAAAG   1620
      GAGAAAGCAG AGGCCTCTAA TGCAGGAGCT GCCTGTATGG GGAACAGCCA GCACCAGGGC   1680
      AGGCAGATGG GGGCGGGGGC ACACCCCCCA ATGATCCTGC CCCTTCCCCT GCGAAAGCCC   1740
      ACCACACTTA GGCAGTGCGA AGTGCTCATC CGCGAGCTGT GGAATACCAA CCTCCTGCAG   1800
      ACCCAAGAGC TGCGGCACCT CAAGTCCCTC CTGGAAGGGA GCCAGAGGCC CCAGGCAGCC   1860
  10  CCGGAGGAAG CTAGCTTTCC CAGGGACCAA GAAGCCACGC ATTTCCCCAA GGTCTCCACC   1920
      AAGAGCCTCT CCAAGAAATG CCTGAGCCCA CCTGTGGCGG AGCGTGCCAT CCTGCCCGCA   1980
      CTGAAGCAGA CCCCGAAGAA CAACTTTGCC GAGAGGCAGA AGAGGCTGCA GGCAATGCAG   2040
      AAACGGCGCC TGCATCGCTC AGTGCTTTGA                                    
```

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      MSGAGVAAGT RPPSSPTPGS RRRRQRPSVG VQSLRPQSPQ LRQSDPQKRN LDLEKSLQFL   60
      QQQBSEMLAK LHEEIEHLKR ENKGEPARGP RPALPPQAHS TLPLPQHRNT AINSSTRLGS   120
      GGTQDGEFLQ TVLAHLAALA PVCQPSGYRF WGTWTDAATS SRGWTMLCSQ AQHVLLSGSP   180
      GPEVIAGRQV ATGCSFDLPP PSRAEMGRNE WDSPCPARSL PQIAAVARPR ISSPMALSPH   240
  25  MLGAQGIWTH SIQGSLPAIW AATMGTRGGS RVLFPCHLSK ALFHPDSGPH PAQDPGLWSQ   300
      AHFPLSLGLG LTSGGHLTGG WSQPGNIAAG AVPRALPSQG DMEKGVEGGP FPSRCGNSSE   360
      LFWAKCGPSR QPQPCSAGDA DRTREEAMLS LGTCCSMCPK PSCFPDGPSG NHLSRASAPL   420
      GARWVCINGV WVEPGGPSPA RLKEGSSRTH RPGGKRGRLA GGSADTVRSP ADSLSMSSFQ   480
      SVKSISNSAN SQGKARPQPG SFNKQDSKAD VSQKADLKEE PLLHNSKLDK VPGVQGQARK   540
      EKAEASNAGA ACMGNSQHQG RQMGAGAHPP MILPLPLRKP TTLRQCEVLI RELWNTNLLQ   600
  30  TQELRHLKSL LEGSQRPQAA PEEASFPRDQ EATHFPKVST KSLSKKCLSP PVAERAILPA   660
      LKQTPKNNFA ERQKRLQAMQ KRRLHRSVL
```

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      GAATTCCGGC CAGAAGAAAT CTGGCCTCGG AACACGCCAT TCTCCGCGCC GCTTCCAATA    60
  40  ACCACTAACA TCCCTAACGA GCATCCGAGC CGAGGGCTCT GCTCGGAAAT CGTCCTGGCC   120
      CAACTCGGCC CTTCGAGCTC TCGAAGATTA CCGCATCTAT TTTTTTTTTC TTTTTTTTCT   180
      TTTCCTAGCG CAGATAAAGT GAGCCCGGAA AGGGAAGGAG GGGGCGGGGA CACCATTGCC   240
      CTGAAAGAAT AAATAAGTAA ATAAACAAAC TGGCTCCTCG CCGCAGCTGG ACGCGGTCGG   300
      TTGAGTCCAG GTTGGGTCGG ACCTGAACCC CTAAAAGCGG AACCGCCTCC CGCCCTCGCC   360
  45  ATCCCGGAGC TGAGTCGCCG GCGGCGGTGG CTGCTGCCAG ACCCGGAGTT TCCTCTTTCA   420
      CTGGATGGAG CTGAACTTTG GGCGGCCAGA GCAGCACAGC TGTCCGGGGA TCGCTGCACG   480
      CTGAGCTTCC TCGGCAAGAC CCAGCGGCGG CTCGGGATTT TTTTGGGGGG GCGGGGACCA   540
      GCCCCGCGCC GGCACCATGT TCCTGGCGAC CCTGTACTTC GCGCTGCCGC TCTTGGACTT   600
      GCTCCTGTCG GCCGAAGTGA GCGGCGGAGA CCGCCTGGAT TGCGTGAAAG CCAGTGATCA   660
  50  GTGCCTGAAG GAGCAGAGCT GCAGCACCAA GTACCGCACG CTAAGGCAGT GGGTGGCCGG   720
      CAAGGAGACC AACTTCAGCC TGGCATCCGG CCTGGAGGCC AAGGATGAGT GCCGCAGCCGC   780
      CATGGAGGCC CTGAAGCAGA AGTCGCTCTA CAACTGCCGC TGCAAGCGGG GTATGAAGAA   840
      GGAGAAGAAC TGCCTGCGCA TTTACTGGAG CATGTACCAG AGCCTGCAGG GAAATGATCT   900
      GCTGGAGGAT TCCCCATATG AACCAGTTAA CAGCAGATTG TCAGATATAT TCCGGGTGGT   960
  55  CCCATTCATA TCAGATGTTT TTCAGCAAGT GGAGCACATT CCCAAAGGGA CAAACTGCCT   1020
      GGATGCAGCG AAGGCCTGCA ACCTCGACGA CATTTGCAAG AAGTACAGGT CGGCGTACCAT   1080
      CACCCCGTGC ACCACCAGCG TGTCCAACGA TGTCTGCAAC CGCCGCAAGT GCCACAAGGC   1140
      CCTCCGGCAG TTCTTTGACA AGGTCCCGGC CAAGCACAGC TACGGAATGC TCTTCTGCTC   1200
      CTGCCGGGAC ATCGCCTGCA CAGAGCGGAG GCGACAGACC ATCGTGCCTG TGTGCTCCTA   1260
  60  TGAAGAGAGG GAGAAGCCCA ACTGTTTGAA TTTGCAGGAC TCCTGCAAGA CGAATTACAT   1320
      CTGCAGATCT CGCCTTGCGG ATTTTTTTAC CAACTGCCAG CCAGAGTCAA GGTCTGTCAG   1380
      CAGCTGTCTA AAGGAAAACT ACGCTGACTG CCTCCTCGCC TACTCGGGGC TTATTGGCAC   1440
      AGTCATGACC CCCAACTACA TAGACTCCAG TAGCCTCAGT GTGGCCCCAT GGTGTGACTG   1500
      CAGCAACAGT GGGAACGACC TAGAAGAGTG CTTGAAATTT TTGAATTTCT TCAAGGACAA   1560
  65  TACATGTCTT AAAAATGCAA TTCAAGCCTT TGGCAATGGC TCCGATGTGA CCGGTGTGGCA   1620
      GCCAGCCTTC CCAGTACAGA CCACCACTGC CACTACCAA  ACTGCCCTCC GGGTTAAGAA   1680
      CAAGCCCCTG GGGCCAGCAG GGTCTGAGAA TGAAATTCCC ACTCATGTTT TGCCACCGTG   1740
      TGCAAATTTA CAGGCACAGA AGCTGAAATC CAATGTGTCG GGCAATACAC ACCTCTGTAT   1800
      TTCCAATGGT AATTATGAAA AAGAAGGTCT CGGTGCTTCA AGCCACATAA CCACAAAATC   1860
  70  AATGGCTGCT CCTCCAAGCT GTGGTCTGAG CCCACTGCTG GTCCTGGTGG TAACCGCTCT   1920
      GTCCACCCTA TTATCTTTAA CAGAAACATC ATAGCTGCAT TAAAAAAATA CAATATGGAC   1980
      ATGTAAAAAG ACAAAAACCA AGTTATCTGT TTCCTGTTCT CTTGTATAGC TGAAATTCCA   2040
      GTTTAGGAGC TCAGTTGAGA AACAGTTCCA TTCAACTGGA ACATTTTTTT TTTTCCTTTT   2100
      AAGAAAGCTT CTTGTGATCC TTCGGGGCTT CTGTGAAAAA CCTGATGCAG TGCTCCATCC   2160
  75  AAACTCAGAA GGCTTGGGA  TATGCTGTAT TTTAAAGGGA CAGTTTGTAA CTTGGGCTGT   2220
      AAAGCAAACT GGGGCTGTGT TTTCGATGAT GATGATCATC ATGATCATGA TGATTTTAAC   2280
      AGTTTTACTT CTGGCCTTTC CTAGCTAGAG AAGGAGTTAA TATTTCTAAG GTAACTCCCA   2340
      TATCTCCTTT AATGACATTG ATTTCTAATG ATATAAATTT CAGCCTACAT TGATGCCAAG   2400
      CTTTTTTGCC ACAAAGAAGA TTCTTACCAA GAGTGGGCTT TGTGGAAACA GCTGGTACTG   2460
  80  ATGTTCACCT TTATATATGT ACTAGCATTT TCCACGCTGA TGTTTATGTA CTGTAAACAG   2520
      TTCTGCACTC TTGTACAAAA GAAAAAACCA CCCGGAATTC
```

982

```
        1         11        21        31        41        51
        |         |         |         |         |         |
     MFLATLYFAL PLLDLLLSAE VSGGDRLDCV KASDQCLKEQ SCSTKYRTLR QCVAGKETNF    60
     SLASGLEAKD SCRSAMEALK QKSLYNCRCK RGMKKEKNCL RIYWSMYQSL QGNDLLEDSP   120
     YEPVNSRLSD IFRVVPFISD VPQQVEHIPK GNNCLDAAKA CNLDDICKKY RSAYITPCTT   180
     SVSNDVCNRR KCHKALRQFF DKVPAKHSYG MLFCSCRDIA CTERRRQTIV PVCSYEEREK   240
     PNCLNLQDSC KTNYICRSRL ADFFTNCQPE SRSVSSCLKE NYADCLLAYS GLIGTVMTPN   300
     YIDSSSLSVA PWCDCSNSGN DLEECLKFLN FFKDNTCLKN AIQAPGNGSD VTVWQPAPPV   360
     QTTTATTTTA LRVKNKPLGP AGSENEIPTH VLPPCANLQA QKLKSNVSGN THLCISNGNY   420
     EKEGLGASSH ITTKSMAAPP SCGLSPLLVL VVTALSTLLS LTETS
```

```
Seq ID NO: 85 DNA sequence
Nucleic Acid Accession #: XM_027172.1
Coding sequence: 143..1405
```

```
        1         11        21        31        41        51
        |         |         |         |         |         |
     GGTGGAGACA CCGCCTCAGG GCTCGGTGCA CAGTGGACAT TTGGGGAGCG TTGTGGGTGA    60
     CCCCCACACA GGCACTGGGA ATGCAGGGGA GAGGGGGCCA AGGGGGAAAG GGGCCAGAGT   120
     GTTGGCTTTG GATTCAGGAG GGATGGATTC CAGTCCTAGC TTGCCACTTA TTAGGACTCC   180
     TGAGAGCAGC CTCCATGAGG CCCTGGACCA GTGCATGACC GCCCTGGACC TCTTCCTCAC   240
     CAACCAGTTC TCAGAAGCAC TCAGCTACCT CAAGCCCAGA ACCAAGGAAA GCATGTACCA   300
     CTCACTGACA TATGCCACCA TCCTGGAGAT GCAGGCCATG ATGACCTTTG ACCCTCAGGA   360
     CATCCTGCTT GCCGGCAACA TGATGAAGGA GGCACAGATG CTGTGTCAGA GGCACCGGAG   420
     GAAGTCTTCT GTAACAGATT CCTTCAGCAG CCTGGTGAAC CGCCCCACGC TGGGCCAATT   480
     CACTGAAGAG GAAATCCACG CTGAGGTCTG CTATGCGAGA TGCCTGCTGC AGCGAGCAGC   540
     CCTGACCTTC CTGCAGGGTT CCTCACACGG AGGGGCAGTC AGGCCCAGAG CCTTGCATGA   600
     TCCCTCTCAC GCCTGCAGCT GCCCACCTGG GCCAGGCCGT CAGCATCTTT TCCTCCTGCA   660
     GGACGGACA ATGGTGAGCT TCATCAAAGG CGGCATCAAA GTTCGAAACA GCTACCAGAC   720
     CTACAAGGAG CTGGACAGCC TTGTTCAGTC CTCACAAATAC TGCAAGGGTG AGAACCACCC   780
     GCACTTTGAA GGAGGAGTGA AGCTTGGTGT AGGGGCCTTC AACCTGACAC TGTCCATGCT   840
     TCCTACTAGG ATCCTGAGGC TGTTGGAGTT TGTGGGGTTT TCAGGAAACA AGGACTATGG   900
     GCTGCTGCAG CTGGACGAGG GAGCGGTCAGG GCACAGCTTC CGCTCTGTGC TCTGTGTCAT   960
     GCTCCTGCTG TGCTACCACA CCTTCCTCAC CTTCGTGCTC GGTACTGGGA ACGTCAACAT  1020
     CGAGGAGGCC GAGAAGCTCT TGAAGCCCTA CCTGAACCGG TACCCTAAGG GTGCCATCTT  1080
     CCTGTTCTTT GCAGGGAGGA TTGAAGTCAT TAAAGGCAAC ATTGATGCAG TGAGTGATGG  1140
     GGGTCCGGGC CGGGGCTGGG GATCCCTCGG GGTCTCCCAG ACCAGCAGGA AGTCAGGCAC  1200
     ATGTGACATA CTCAGGGACA GGATAGACTG GGGGCTGGGG GGGGGCCAAG AGAGAACCAA  1260
     CCAGAGAGCA GGGGCAGGAG AGGCCCTTCT GGCAGGAAGA CAAGGGAGGA  1320
     GGAGGCCATT GTGGTGCCTG GGATTTTGAC TGGGAGATAT AGGACTGCAG CATTGCAGTG  1380
     GAGGGAGGTG GAGGGAGGTG CTTGAGGGAG GCAGAGGTTA GGAAAGCCCA TCTGTTTAGG  1440
     GCATGACGAT TAGGCTGGAG TCTGGTACCT CCCCTCCATT ATAGCTCTCT CCTGCTCTTT  1500
     CATTTTGTTA CTAAAAACCA GAGTCCTAGG CGGGGGCTGT ATTTGAGCCC AACGTCATGT  1560
     AAGACTTAGG AGGTAAAACC AGGACTGGAG GCCAGATCTC CTGGCTCCTG GGGCCCCACC  1620
     TGAGCCTAGC ACAGGGCTGG ACCACTATGC CCTGGAGGAG TCCCGGTCTG CTGTGGTCTT  1680
     GGGAGGTTCG GAGGATGCAG AGGGGTTGGG GCTGGGTGGG CACCCGGTCAG GCTGACCAGA  1740
     AGGTGCCTGC AGGCCATCCG GCGTTTCGAG GAGTGCTGTG AGGCCCAGCA GCACTGGAAG  1800
     CAGTTCCACC ACATGTGCTA CTGGGAGCTG ATGTGGTGCT TCACCTACAA GGGCCAGTGG  1860
     AAGATGTCCT ACTTCTACGC CGACCTGCTC AGCAAGGAGA ACTGCTGGTC CAAGGTGGGC  1920
     TGATGCCACG TGTTAGGGGC ATTGGGTGAC CAGGGCTGAC TGTGTGCCTC CAGACCACGG  1980
     GCCAAATCCC TAACTGAACA CAGATGTCTC AGCTGGAATC TAAACATAAC CTTAAATTCT  2040
     AATAGGACTC AGGCTTGGAA GGAACTAAAG ACCACAAGAG AAACTTCTGA ACCACAAATGT  2100
     GTCACAAAGA GATTTCTAGT CACAAAGGAC AGAAACATGG CTCCCTCTGT CCAGTAGAAC  2160
     TGTTTCTGTG GTAGAAATGT TCTACTTATG CACTGCCCAA ATATGGTAGT CACCCAGCCAC  2220
     ATGTGCACAA TGAGCACATG AAATGTGCCT AGTGCAACTG GGAAACTGAT TGTTTTCAGT  2280
     TTTATTTAAT TTTAATTAAC TAAATGTTAA ATTTAAATAG CCATGTAGGG CTTGTGGCCA  2340
     CTATATTGGA CTATGCAGGT CCAAAACACA AAAGGCTCAT ATAACTGAAC ATTCTTGGCA  2400
     CATCCGACTT CAGGTAGGGC TGGATCCAGG AATTCAAATG ATGTCGTCTG GCTTGGTCTT  2460
     TCCATTTGTG GCTGTGCTCT CTCCTATGAC ATCTTGTTT CTGCTGCATC TTTGCAGGAA  2520
     GGTTCTCTCC ATGTGACAGG CAAGGTGGCC ACGGGCTGCT TCTACTCATA TCCTCCCCTT  2580
     GGTTTCAACC AGAGTCCCAC AGTTTGAGCC TCACTGGTCAC CTGCCAATGC  2640
     CTGGAACAGA GGGGTGGGAA GATTCCTCAT GGGCTGAGAG TAGGAGAGGG GTGGTTCCAC  2700
     AGAAGAAAAT GATGCACCCA GAAATAGTGG AGGGATTAAC AAGATGCCAT ACAGGCAAAA  2760
     CAAAGCCAAA CAGATGCCCG CCTACCAAAG ATGAAATTTA TCATGGTAAG TATTGAAATA  2820
     AGTGTTAGCT TGTACCATCA TAGTAATGAT AGTGCAGAAA TTGGAACCAA GAGTCTTACA  2880
     ACCACCTAGC TCAGCAAACG TCTAATCTGT TATTTGTAAA TACACAGGAC ATGTGTCTTC  2940
     ATGGCTTCAT CCCCTGCCCA TAGCAGACAT TGCTAATCAA TCCTCTGCCA TGAGCCTAGC  3000
     TGTGACCTTA GATACCTTCC CTGCAAAGCT CCAGGCAGCC GTTAGAACTG ACCCACATTG  3060
     GCACCTGTTA GGAGACCTGT CATCTGCATC TGGTCCCTCC CCCTCCTCTG CATGGAGGTC  3120
     TCTTCCGTAG CCCTGCCTGG GGACCAAGCT GAAGGGGAAT TCTCTGGGTC TGGGCAGGT  3180
     GGGGCTGGGG AAGGGAGCAG GCTTCTACTG AGCTCTCAAC CCTGTCATGT  3240
     CCTCCACACC ACACCCAGTG AGGAGAAGTT ATTATCTCCA TTGGGCACAT GAGGGCCACT  3300
     GATGCTAAGC GAGGACTGTG ACTTTCCCAC CCACATGGCT GTTGGTGACA GACACCGGAT  3360
     TTGAGGACAG TCCTGTGTGA CTCAGAGGCC TCTGCTTCTC TCTTGCTCCT AGAGTCCTCC  3420
     TGAGGAGTCG GGGCTTGCCC TGAGCCCACC CTGCTGTTGA AGGTGCTTCC TCAGGCCCAG  3480
     CTCCCATGGC CCCCACACCC CCTCCTCATC ACCTCCTACT CCCAAAAAGG ACAAAGCCTC  3540
     AGGGAACCTT TTTTCTTTTT TTAGAGACAG GGTCTTGCTA TGTTGGTCAG GCTAGTCTTG  3600
     AACTCCTGGG CCCAAGCAAT CTTCCCACCT CTGCCTCCTA AAGTGCTGGG ATTCTGGGCC  3660
     TTAGGGAACC TTTTTGAACT GAAAGTGACT CTCGAAGCCT TCCTGTAGAG TGAGGTGGGT  3720
     GGCTGGGGCT AACCATATGG GGAAGGAGAG ACCTGGTGGG GGCACACAGC TGCTATATAG  3780
     AGGAACAGAA GGTGGGCCAG GCCTCCCAGC TGTCAGGAAA TCTGGCTCAG TCCCCAGACC  3840
     GCTGTGTGGC TTTGGGGTGG CCAGTCCCTC TCTCTCTGCA ACACTCTCCT CATGAGGACT  3900
     TCTGTGAAAA ATGGGGGTTG TAACCCTCTG AGTTCTGGGG CCCCACCCAT TTCTAATCTC  3960
```

```
CAGAACTTGG CCACCACAGA CTCCACCAAC TTCTAGTCCT GGGGCCTGGG CCTCTGGCCA    4020
TTGCCATAGG CACCACCTGC TCTGTGCAGG CAGCGCCCCC CTCTGCCAGG ATCCTCCGAG    4080
GTCAGCTGCT GGGTCTGACC CGCAGACCCT GGCTGAGCGA CGGATGAACG GAGTATGCAG    4140
ACACAGGCTT GGCCTGTCAGC AGATGGGGGA CCCCTGCCAG AGTCAGCAGC GGCCCCCATA   4200
AGCCTGCCAC GCTTGCATTT ATTTAGTACA GATGTAATGA CAAAGGCCTA AAGCAAACTC    4260
CATTTGTGGG TAATTAACAT TGTCGCCCCC CCAGAAAGAG CAGTCCTCCG CATGATGATT     4320
AAAGGCCAGG TTCCGAGGCC TAAGTAAACC AACTTATCTA GATCAATTCC CTTACTTCTT     4380
GTTATCTACT CTGAGAGAAT TCAGCTGCCT TCAGCCAAAT CCTTTCCCGA AGCTTTTGCA     4440
AAACCTCCGA GCCTTCCAAG GTTTGCTTCT TTCTGTAATT TTTCTCACCA CCCTGACCTA     4500
TCTCCTGCAG TCAGCCCTGT GGAGGCCTTT GTGTTTCCCC CAGTGCTGGC AGCCTAGAGG     4560
CTGAGATGGC CAGAAACAAG GTGGTGACAG TGGCGTGCTC AGGGCTTGGG AAACCCAAGG     4620
AGCTAAAGGC ATGCCCAGGC AACCAAAGAG GACAGGAAGG CTTCTGAGGA GAGACCTCTG     4680
AGGTGGGTCT TGGAGAGGAA GGACTTAGGG AGGCAGAGTG GAGGAAGTGA GAGGACACCC     4740
CAAGCCAAGA GGGCGGCAGG ACCAAAGGCT CAGAAGCCAG GGCGCTGCAG AGGGGCTGTG     4800
TGCCACAGGG TGAAGAGTTT GTGTGGCAGA AGGGCAGGGG GCTTGCATCA GGGGTGACAG     4860
CTGCTCTTTT GTCCCAGCAT AGCCCCTGTA CATCCCTGGA GAGCTGGGGC GTCCGCAACT     4920
CTAAGTCACA GCCCCCATCC TAACCCTGGT GGTGCAGTGA GGGTGAGCTG TCTGTGGGCA     4980
GGAGGGAAGA CTCTTGGAGA TGAGCCTGGT GAAGGGATAA TGGCATCCCG GGCCGAGGAG     5040
CAGCACAGGC AGAGGCCTGG GGAGAGTTTA AGGAGTGTAG GGGAGGAAAT GGCAGAAGAT     5100
GAGCCAGAAA AAGAAAGGTT AGGGCAGGTC CTGGAGGACA TGAGTGGCTG TTTGGGCTTT     5160
ATCCAGCAGT GGGGGAGCCT TGGCAGGCTT GTGGCTTAGA TAGGTGCTTT AGAAAGCCCA     5220
CCAGCAGTTG CTGGGCCACC CCGCTGGCTG GGTCCTGTTC TAAGGCAGGA AATACAAGCA     5280
TGAGCAGGAA AAGACCCCCT CAAGGCTCAC GTCCTAGTGG GGAGACAAGA AACACAGATG     5340
GGCAATATAA CACGATGTCT GGTTCCAGTA AGTGCAGTGA AGAACAAGCG AGGCTGGATG     5400
CAGGGAGTGA TGGGAGGGGC TTTGTAAGGG GAGGTCGGGG GAAGCCTGTC TCAGAGGACA     5460
CCAGAATGGA GCGCAGGGAGC AGCATGTGGC AGTCACATGG CAGGCCGTTA GGGCAGAGGG    5520
AGCTCGGGCAG GGCACAGCAG GGCAGGGAGTG TGTTTGATGT GTCCTGGGAA CCGCCCTGAG   5580
GCCGTCGTGT GGCTGGAGTG CTGCAGGTGT CAAGGAAATT GTAGGAGATG TCTCCTGAGT     5640
GTGATGGAAT ATAACCAGAT TTCCAGAAGG AACTGACATG ATCTGACTTA AAAAGGTCAG     5700
TGTGCGAAAT GGCTTGCAGG GGACAGGAGT GGGAGCAGGG AGATAGGAGA CAATGTGTAC     5760
CAGGACAGCA GAAAGACATC CCGGGTAGCC TGGAACAGGG AGACGGTGTG GAGATGGTGG     5820
CAGTCCGATA ATGAGAGCCG TAGGGCAAGG CCAGCAGGAT CCTAGAGTGA GACGGGAGGT     5880
AAAGTCACCG GGACTTGGTG TCTCCACGTC AGGGGCAGGG GAAAGGGAGA GGACAAGGGT     5940
GACCCGGGAG GTTAAAGATG GGACCGGGATG CAGACGCAGT GGCTCATGCC TGTAATCCTA    6000
GCACTTTGGG AGGCTGAGGC GGGCGGATAG CTTGAGGTCA GGAGTTTGAA ACCAGCCTGG     6060
CCAACATGGT GAAACCCCGT CTCTACTAAA ATATACAAAA ATTAGCCTGG CGTGGTGGTG     6120
CATGCCTGCA GTCCCAGCTA TTCAGGAGGC TGAGGCAACA AGAATCGCTT GAACCTGGGA     6180
GGCGGAGGTT GCAGTGAGCC GAGATCGCGC CATAGCACTC CAGCCTTAGC CTGGGCGACA     6240
GAGCGAGACC ACATC
```

Seq ID NO: 86 <u>Protein sequence</u>
Protein Accession #: XP_027172.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MDSSPSLPLI RTPESSLHEA LDQCMTALDL FLTNQFSEAL SYLKPRTKES MYESLTYATI    60
LEMQAMMTFD PQDILLAGNM MKEAQMLCQR HRRKSSVTDS PSSLVNRPTL GQFTEEEIHA   120
EVCYAECLLQ RAALTFLQGS SHGGAVRPRA LHDPSHACSC PPGPGRQHLF LLQDENMVSF   180
IKGGIKVRNS YQTYKELDSL VQSSQYCKGE NHPBFEGGVK LGVGAFNLTL SMLPTRILRL   240
LEFVGFSGNK DYGILQLEEG ASGHSFR8VL CVMLLLCYHT FLTFVLGTGN VNIEEAEKLL   300
KPYLNRYPKG AIFLFPAGRI FVIKGNIDAV SDGGPGRGWG SLGVSQTSRK SGTCDILRDR   360
IDWGRGGGQE RTNQRAGAGE ALLAEQPGKT RBEEAFVVPG ILTGRYRTAA LQWREVEGGA
```

Seq ID NO: 87 <u>DNA sequence</u>
Nucleic Acid Accession #: AB007921
Coding sequence: 143..1363

```
1          11         21         31         41         51
|          |          |          |          |          |
GGTGGAGACA CCGCCTCAGG GCTCGGTGCA CAGTGGACAT TTGGGGAGCG TTGTGGGTGA    60
CCCCCACACA GGCACTGGGA ATGCAGGGGA GAGGGGGCCA AGGGGGAAAG GGGCCAGAGT    120
GTTGGCTTTG GATTCAGGAG GGATGGATTC CAGTCCTAGC TTGCCACTTA TTAGGACTCC    180
TGAGAGCAGC CTCCATGAGG CCCTGGACCA GTGCATGACC GCCCTGGACC TCTTCCTCAC    240
CAACCAGTTC TCAGAAGCAC TCAGCTACCT CAAGCCCAGA ACCAAGGAAA GCATGTACCA    300
CTCACTGACA TATGCCACCA TCCTGGAGAT GCAGGCCATG ATGACCTTTG ACCCTCAGGA    360
CATCCTGCTT GCCGGCAACA TGATGAAGGA GGCACAGATG CTGTGTCAGA GGCACCGGAG    420
GAAGTCTTCT GTAACAGATT CCTTCAGCAG CCTGGTGAAC CGCCCCACGC TGGGCCAATT    480
CACTGAAGAG GAAATCCACG CTGAGGTCTG CTATGCAGAG TGCCTGCTGC AGCGAGCGAC    540
CCTGACCTTC CTGCAGGGTT CCTCACACGG AGGGGCAGTC AGGCCCAGAG CCTTGCATGA    600
TCCCTCTCAC GCCTGCAGCT GCCCACCTGG GCCAGGCCGT CAGCATCTTT TCCTCCTGCA    660
GGACGAGAAC ATGGTGAGCT TCATCAAAGG CGGCATCAAA GTTCGAAACA GCTACCAGAC    720
CTACAAGGAG CTGGACAGCC TTGTTCAGTC CTCACAATAC TGCAAGGGTG AGAACCACCC    780
GCACTTTGAA GGAGGAGTGA AGCTTGGTGT AGGGGCCTTC AACCTGACAC TGTCCATGCT    840
TCCTACTAGG ATCCTGAGGC TGTTGGAGTT TGTGGGGTTT TCAGGAAACA AGGACTATGG    900
GCTGCTGCAG CTGGAGGAGG GAGCGTCAGG GCACAGCTTC CGCTCTGTGC TCTGTGTCAT    960
GCTCCTGCTG TGCTACCACA CCTTCCTCAC CTTCGTGCTC GGTACTGGGA ACGTCAACAT   1020
CGAGGAGGCC GAGAAGCTCT TGAAGCCCTA CCTGAACCGG TACCCTAAGG GTGCCATCTT   1080
CCTGTTCTTT GCAGGGAGGA TTGAAGTCAT TAAAGGCAAC ATTGATGCAG TGAGTGATGG   1140
GGGTCCGGGC CGGGGCTGGG GATCCCTCGG GGTCTCCCAG ACCAGCAGGA AGTCAGGCAC   1200
ATGTGACATA CTCAGGGACA GGATAGACTG GGGGCGGGGG GGGGGCCAA GAGAGAACCA    1260
ACCAGAGAGC AGGGGCAGGA GAGGCCCTTC TGGCAGAGCA GCCTGGGAAG ACAAGGGAGG   1320
AGGAGGCATT TGTGGTGCCT GGGATTTTGA CTGGGAGATA TAGGACTGCA GCATTGCAGT   1380
GGAGGGAGGT GGAGGGAGGT GCTTGAGGGA GGCAGAGGTT AGGAAAGCCC ATCTGTTTAG   1440
GGCATGACGA TTAGGCTGGA GTCTGGTACC TCCCCTCCAT TATAGCTCTC TCCTGCTCTT   1500
```

```
     TCATTTTGTT ACTAAAAACC AGAGTCCTAG GCGGGGGCTG TATTTGAGCC CAACGTCATG  1560
     TAAGACTTAG GAGGTAAAAC CAGGACTGGA GGCCAGATCT CCTGGCTCCT GGGGCCCCAC  1620
     CTGAGCCTAG CACAGGGCTG GACCACTATG CCCTGGAGGA GTCCCGGTCT GCTGTGGTGT  1680
     TGGGAGGTTC GGAGGATGCA GAGGGGTTGG GGCTGGGTGG GCACCCGTCA GGCTGACCAG  1740
  5  AAGGTGCCTG CAGGCCATCC GGCGTTTCGA GGAGTGCTGT GAGGCCCAGC AGCACTGGAA  1800
     GCAGTTCCAC CACATGTGCT ACTGGGAGCT GATGTGGTGC TTCACCTACA AGGGCCAGTG  1860
     GAAGATGTCC TACTTCTACG CCGACCTGCT CAGCAAGGAG AACTGCTGGT CCAAGGTGGG  1920
     CTGATGCCAC GTGGTAGGGG CATTGGGTGA CCAGGGCTGA CTGTGTGCCT CCAGACCACG  1980
     GGCCAAATCC CTAACTGAAC ACAGATGTCT CAGCTGGAAT CTAAACATAA CCTTAAATTC  2040
 10  TAATAGGACT CAGGCTTGGA AGGAACTAAA GACCACAAGA GAAACTTCTG AACCACAATG  2100
     TGTCACAAAG AGATTTCTAG TCACAAAGGA CAGAAACATG GCTCCCTCTG TCCAGTAGAA  2160
     CTGTTTCTGT GGTAGAAATG TTCTACTTAT GCACTGCCCA AATATGGTAG TCACCAGCCA  2220
     CATGTGCACA CTGAGCACAT GAAATGTGCC TAGTGCAACT GGGAAACTGA TTGTTTTCAG  2280
     TTTTATTTAA TTTTAATTAA CTAAATGTTA AATTTAAATA GCCATGTAGG GCTTGTGGCC  2340
 15  ACTATATTGG ACTATGCAGG TCCAAAACAC AAAAGGCTCA TATAACTGAA CATTCTTGGC  2400
     ACATCCGACT TCAGGTAGGG CTGGATCCAG GAATTCAAAT GATGTCGTCT GGCTTGGTCT  2460
     TTCCATTTGT GGCTGTGCTC TCTCCTATGA CATCTTTGTT TCTGCTGCAT CTTTGCAGGA  2520
     AGGTTCTCTC CATGTGACAG GCAAGGTGGC CACGGGCTGC TTCTACTCAT ATCCTCCCCT  2580
     TGGTTTCAAC CAGAGTCCCA CAGTTTGAGC CTCACTGGTC TGACTTGTCA CCTGCCAATG  2640
 20  CCTGGAACAG AGGGGTGGGA AGATTCCTCA TGGGCTGAGA GTAGGAGAGG GGTGGTTCCA  2700
     CAGAAGAAAA TGATGCACCC AGAAATAGTG GAGGGATTAA CAAGATGCCA TACAGGCAAA  2760
     ACAAAGCCAA ACAGATGCCC GCCTACCAAA GATGAAATTT ATCATGGTAA GTATTGAAAT  2820
     AAGTGTTAGC TTGTACCATC ATAGTAATGA TAGTGCAGAA ATTGGAACCA AGAGTCTTAC  2880
     AACCACCTAG CTCAGCAAAC GTCTAATCTG TTATTTGTAA ATACACAGGA CATGTGTCTT  2940
 25  CATGGCTTCA TCCCCTGCCC ATGGCAGACA TTGCTAATCA ATCCTCTGCC ATGAGCCTAG  3000
     CTGTGACCTT AGATACCTTC CCTGCAAAGC TCCAGGCAGC CGTTAGAACT GACCCACATT  3060
     GGCACCTGTT AGGAGACCTG TCATCTGCAT CTGGTCCCTC CCCCTCCTCT GCATGGGAGGT  3120
     CTCTTCCGTA GCCCTGCCTG GGGACCAAGC TGAAGGGGAA TTCTCTGGGT CTGGGGCAGG  3180
     TGGGGCTGGG GAAGGGAGCA GGCTTCTACT GAGCTCTCAA CATGTCTGGC ACCTGTCATG  3240
 30  TCCTCCACAC CACACCCAGT GAGGAGAAGT TATTATCTCC ATTGGGCACA TGAGGGCCAC  3300
     TGATGCTAAG CGAGGACTGT GACTTTCCCA CCCACATGGC TGTTGGTGAC AGACACCGGA  3360
     TTTGAGGACA GTCCTGTGTG ACTCAGAGGC CTCTGCTTCT CTCTTGCTCC TAGAGTCCTC  3420
     CTGAGGAGTC GGGGCTTGCC CTGAGCCCAC CCTGCTGTTG AAGGTGCTTC CTCAGGCCCA  3480
     GCTCCCATGG CCCCCACACC CCCTCCTCAT CACCTCCTAC TCCCAAAAAG GACAAAGCCT  3540
 35  CAGGGAACCT TTTTTCTTTT TTTAGAGACG GGGTCTTGCT ATGTTGGTCA GGCTGGTCTT  3600
     GAACTCCTGG GCCCAGGCAA TCTTCCCGCC TCTGCCTCCT AAAGTGCTGG GATTCTGGGC  3660
     CTTAGGGAAC CTTTTTGAAC TGAAAGTGAC GCTCGAACCC TTCCTGTAGA GTGAGGTGGG  3720
     TGGCTGGGGC TAACCATATG GGGAAGGAGA GACCTGGTGG GGGCACACAG CTGCTATATA  3780
     GAGGAACAGA AGGTGGGCCA GGCCTCCCAG CTGTCAGGAA ATCTGGCTCA GTCCCCAGAC  3840
 40  CGCTGTGTGG CTTTGGGGTG GCCAGTCCCT CTCTCTCTGC AACACTCTTC TCATGGAGGAC  3900
     TTCTGTGAAA AATGGGGGTT GTAACCCTCT GAGTTCTGGG GCCCCACCCA TTTCTAATCT  3960
     CCAGAACTTG GCCACCACAG ACTCCACCAA CTTCTAGTCC TGGGGCCTGG GCCTCTGGCC  4020
     ATTGCCATAG GCACCACCTG CTCTGTGCAG GCAGCGCCCC CCTCTGCCAG GATCCTCCGA  4080
     GGTCAGCTGC TGGGTCTGAC CCGCAGACCC TGGCTGGACG ACGGATGAAC GGAGTATGCA  4140
 45  GACACAGGCT TGCTTGTCAG CAGATGGGGG ACCCCTGCCA GAGTCAGCAG CGGCCCCCAT  4200
     AAGCCTGCCA CGCTTGCATT TATTTAGTAC AGATGTAATG ACAAAGGCCT AAAGCAAACT  4260
     CCATTGTGG GTAATTAACA TTGTCGCCCC CCCAGAAAGA GCAGTCCTCC GCATGATGAT  4320
     TAAAGGCCAG GTTCCGAGGC CTAAGTAAAC CAACTTATCT AGATCAATTC CCTTACTTCT  4380
     TGTTATCTAC TCTGAGAGAA TTCAGCTGCC TTCAGCCAAA TCCTTTCCCG AAGCTTTTGC  4440
 50  AAAACCTCCG AGCCTTCCAA GGTTTGCTTC TTTCTGTAAT TTTTCTCACC ACCCTGACCT  4500
     ATCTCCTGCA GTCAGCCCTG TGGAGGGCCTT TGTGTTTCCC CCAGTGCTGG CAGCCTAGAG  4560
     GCTGAGATGG CCAGAAACAA GGTGGTGACA GTGGCGTGCT CAGGGCTTGG GAAACCCAAG  4620
     GAGCTAAAGG CATGCCCAGG CAACCAAAGA GGACAGGAAG GCTTCTGAGG AGAGACCTCT  4680
     GAGGTGGGTC TTGGAGAGGA AGGACTTAGG GAGGCAGAGT GGAGGAAGTG AGAGGACACC  4740
 55  CCAAGCCAAG AGGGCGGCAG GACCAAAGGC TCAGAAGCCA GGGCGCTGCA GAGGGGCTGT  4800
     GTGCCACAGG GTGAAGAGTT TGTGTGGCAG AAGGGCAGGG GGCTTGCATC AGGGGGTGACA  4860
     GCTGCTCTTT TGTCCCAGCA TAGCCCCTGT ACATCCCTGG AGAGCTGGGG CGTCCACAAC  4920
     TCTAAGTCAC AGCCCCCATC CTAACCCTGG TGGTGCAGTG AGGGTGAGCT GTCTGTGGGC  4980
     AGGAGGGAGA ACTCTTGGAG ATGAGCCTGG TGAAGGGATA ATGGCATCCC GGGCCGAGGA  5040
 60  GCAGCACAGG CAGAGGCCTG GGGAGAGTTT AAGGAGTGTA GGGGAGGAAA TGGCAGAAGA  5100
     TGAGCCAGAA AAAGAAAGGT TAGGGCAGGT CCTGGAGGAC ATGAGTGGCT GTTTGGGCTT  5160
     TATCCAGCAG TGGGGGAGCC TTGGCAGGCT TGTGGCTTAG ATAGGTGCTT TAGAAAGCCC  5220
     ACCAGCAGTT GCTGGGCCAC CCCGCTGGCT GGGTCCTGTT CTAAGGCAGG AAATACAAGA  5280
     ATGAGCAGGA AAAGACCCCC TCAAGGCTCA CGTCCTAGTG GGGAGACAAG AAACACAGAT  5340
 65  GGGCAATATA ACACGATGTC TGGTTCCAGT AAGTGCAGTG AAGAACAAGC GAGGCTGGAT  5400
     GCAGGGAGTG ATGGGAGGGG CTTTGTAAGG GGAGGTCGGG GGAAGCCTGT CTCAGAGGAC  5460
     ACCAGAATGG AGCGCAGGAG CAGCACGTGG CAGTCACATG GCAGGCCGTT AGGGCAGAGG  5520
     GAGCTGGGCA GGGCACAGCA GGGCAGGAGT GTGTTTGATG TGTCCTGGGA ACCGCCCTGA  5580
     GGCCGTCGTG TGGCTGGAGT GCTGCGACGTG TCAAGGAAAT TGTAGGAGAT GTCTCCTGAG  5640
 70  TGTGATGGAA TATAACCAGA TTTCCAGAAG GAACTGACAT GATCTGACTT AAAAAGGTCA  5700
     GTGTGCGAAA TGGCTTGCAG GGGACAGGAG TGGGAGCAGG GAGATAGGAG ACAATGTGTA  5760
     CCAGGACGAC AGAAAGACAT CCCGGGTAGC CTGGAACAGG GAGACGGGTGT GGAGATGGTG  5820
     GCAGTCCGAT AATGAGAGCC GTAGGGCAAG GCCAGCAGGA TCCTAGAGTG AGACGGGAGG  5880
     TAAAGTCACC GGGACTTGGT GTCTCCACGT CAGGGGCAGG GGAAAGGGAG AGGACAAGGG  5940
 75  TGACCCGGGA GGTTAAAGAT GGGACCGGGG CCAGACGCAG TGGCTCATGC CTGTAATCCT  6000
     AGCACTTTGG GAGGCTGAGG CGGGCGGATG GCTTGAGGTC AGGAGTTTGA AACCGGCCTG  6060
     GCCAACATGG TGAAACCCCG TCTCTACTAA AATATACAAA AATTAGCCTG GCGTGGTGGT  6120
     GCATGCCTGC AGTCCCAGCT ATTCAGGAGG CTGAGGCGAC AAGAATCGCT TGAACCTGGG  6180
     AGGCGGAGGT TGCAGTGAGC CGAGATCGCG CCATAGCACT CCAGCCTTAG CCTGGGCGAC  6240
 80  AGAGCGAGAC CACATC
```

Seq ID NO: 88 **Protein sequence**
Protein Accession #: BAA32297.1

```
      1           11          21          31          41          51
      |           |           |           |           |           |
5     VETPPQGSVH  SGHLGSVVGD  PHTGTGNAGE  RGPRGKGARV  LALDSGGMDS  SPSLPLIRTP   60
      ESSLHEALDQ  CMTALDLPLT  NQFSEALSYL  KPRTKESMYH  SLTYATILEM  QAMMTFDPQD  120
      ILLAGNMMKE  AQMLCQRHRR  KSSVTDSFSS  LVNRPTLGQF  TEEEIHAEVC  YAECLLQRAA  180
      LTFLQGSSHG  GAVRPRALHD  PSHACSCPPG  PGRQHLFLLQ  DENMVSPIKG  GIKVRNSYQT  240
      YKELDSLVQS  SQYCKGENHP  HPEGGVKLGV  GAFNLTLSML  PTRILRLLKF  VGFSGNKDYG  300
10    LLQLEEGASG  HSFRSVLCVM  LLLCYHTFLT  FVLGTGNVNI  EEAEKLLKPY  LNRYPKGAIF  360
      LPFAGRIBVI  KGNIDAVSDG  GPGRGWGSLG  VSQTSRKSGT  CDILRDRIDW  GRGGGFRENQ  420
      PESRGRRGPS  GRAAWEDKGG  GGICGAWDFD  WEI
```

Seq ID NO: 89 DNA sequence
Nucleic Acid Accession #: AF007170
Coding sequence: 73..1725

```
      1           11          21          31          41          51
      |           |           |           |           |           |
20    AAGGAGGCGG  CCTCCGGGAA  AAGCGACCGC  AGGACTCCTG  AGAGCAGCCT  CCATGAGGCC   60
      CTGGACCAGT  GCATGACCGC  CCTGGACCTC  TTCCTCACCA  ACCAGTTCTC  AGAAGCACTC  120
      AGCTACCTCA  AGCCCAGAAC  CAAGGAAAGC  ATGTACCACT  CACTGACATA  TGCCACCATC  180
      CTGGAGATGC  AGGCCATGAT  GACCTTTGAC  CCTCAGGACA  TCCTGCTTGC  CGGCAACATG  240
      ATGAAGGAGG  CACAGATGCT  GTGTCAGAGG  CACCGGAGGA  AGTCTTCTGT  AACAGATTCC  300
25    TTCAGCAGCC  TGGTGAACCG  CCCCACGCTG  GGCCAATTCA  CTGAAGAAGA  AATCCACGCT  360
      GAGGTCTGCT  ATGCAGAGTG  CCTGCTGCAG  CGAGCAGCCC  TGACCTTCCT  GCAGGACGAG  420
      AACATGGTGA  GCTTCATCAA  AGGCGGCATC  AAAGTTCGAA  ACAGCTACCA  GACCTACAAG  480
      GAGCTGGACA  GCCTGTTCA   GTCCTCACAA  TACTGCAAGG  GTGAGAACCA  CCCGCACTTT  540
      GAAGGAGGAG  TGAAGCTTGG  TGTAGGGCGT  TTCAACCTGA  CACTGTCCAT  GCTTCCTACT  600
30    AGGATCCTGA  GGCTGTTGGA  GTTTGTGGGG  TTTTCAGGAA  ACAAGGACTA  TGGGCTGCTG  660
      CAGCTGGAGG  AGGGGAGCGTC  AGGGCACAGC  TTCCGCTCTG  TGCTCTGTGT  CATGCTCCTG  720
      CTGTGCTACC  ACACCTTCCT  CACCTTCGTG  CTCGGTACTG  GGAACGTCAA  CATCGAGGAG  780
      GCCGAGAAGC  TCTTGAAGCC  CTACCTGAAC  CGGTACCCTA  AGGGTGCCAT  CTTCCTGTTC  840
      TTTGCAGGGA  GGATTGAAGT  CATTAAAGGC  AACATTGATG  CAGCCATCCG  GCGTTTCGAG  900
35    GAGTGCTGTG  AGGCCCAGCA  GCACTGGAAG  CAGTTTCACC  ACATGTGCTA  CTGGGAGCTG  960
      ATGTGGTGCT  TCACCTACAA  GGGCCAGTGG  AAGATGTCCT  ACTTCTACGC  CGACCTGCTC  1020
      AGCAAGGAGA  ACTGCTGGTC  CAAGGCCACC  TACATTTACA  TGAAGGCCGC  CTACCTCAGC  1080
      ATGTTGGGA   AGGAGGACCA  CAAGCCGTTC  GGGGACGACG  AAGTGGAATT  ATTTCGAGCT  1140
      GTGCCAGGCC  TGAAGCTCAA  GATTGCTGGG  AAATCTCTAC  CCACAGAGAA  GTTTGCCATC  1200
      CGGAAGTCCC  GGCGCTACTT  CTCCTCCAAC  CCTATCTCGC  TGCCAGTGCC  TGCTCTGGAA  1260
40    ATGATGTACA  TCTGGAACGG  CTACGCCGTG  ATTGGGAAGC  AGCCGAAACT  CACGGATGGG  1320
      ATACTTGAGA  TTATCACTAA  GGCTGAAGAG  ATGCTGGAGA  AAGGCCCAGA  GAACGAGTAC  1380
      TCAGTGGATG  ACGAGTGCTT  GGTGAAATTG  TTGAAAGGCC  TGTGTCTGAA  ATACCTGGGC  1440
      CGTGTCCAGG  AGGCCGAGGA  GAATTTTAGG  AGCATCTCTG  CCAATGAAAA  GAAGATTAAA  1500
      TATGACCACT  ACTTGATCCC  AAACGCCCTG  CTGGAGCTGG  CCCTGCTGCT  TATGGAGCAA  1560
45    GACAGAAACG  AAGAGGCCAT  CAAACTTTTG  GAATCTGCCA  AGCAAAACTA  CAAGAATTAC  1620
      TCCATGGAGT  CAAGGACACA  CTTTCGAATC  CAGGCAGCCA  CACTCCAGGC  CAAGTCTTCC  1680
      CTAGAGAACA  GCAGCAGATC  CATGGTCTCA  TCAGTGTCCT  TGTAGCTTTG  TGCAGCAGTT  1740
      CCGGGCTGGA  AGACAGAGAC  AGCTGGACAG  AGCTCCTGAA  AACATTTCAA  AATACCCCCT  1800
      CCCCCTGCCC  TGCCCTGCCT  TTGGGGTCCA  CCGGCACTCC  AGTTGGATGG  CACAACATAG  1860
50    TGTATCCGTG  CAGAAGCCGA  GCTGGCCATTT  TCACCAGTGT  AGCCAAGGGC  CTTTGCCAAG  1920
      GGCAGAGCAG  GTGGAGCCCT  CTGCCTGCCC  TATCACACAT  ACGGGTACTT  GCTTTTCACT  1980
      GTGATGTTTA  AGAGAATGTA  TGAACAGTTT  ACATTTTCCT  TAGAAATACA  TTGATGGGAT  2040
      CACAGTTGGC  TTTAAAACC   AACAACAATC  AACCACCTGT  AAGTCTTTGT  CTTCACCTAT  2100
      TATCATCTGG  AGGTAAATCT  CTTTTATATGA TGATGCCAAA  GGGCAAATTG  CTTTTCAAAT  2160
55    TCAGCAAGTT  CTCAGCTTGT  GTGACGGAAG  GTCCTTCAGA  GGACCTGAGG  AATGCCTGGG  2220
      AGAGGCTAGG  CCTCAGGCTT  CAATGCTTCT  GGGGTTGGGC  ATGAGGATGT  ACACGACAC   2280
      CCACTACCTT  ACTACTCACA  CTTCATTTCA  CTCCTTTTGT  AAATTTCCAA  TTTAAAAATC  2340
      AAGCACGTCT  TTTTAGTGAG  ATAAAATCTG  AGCTCTTCTG  TAGAAAAATC  AATCTCTACC  2400
      AGTAGAAAAT  GCCAGGGCTT  GATGGAAGAG  CTGTGTAGCC  CTTCTATGC   CAAAGCCAGG  2460
60    AAATTTGGGG  GGCAGGAGGA  GGTTCTCAGA  ATCCAGTCTG  TATCTTTGCT  GTATGCCAAA  2520
      CTGAAACCAC  TGGGAATAAT  TTATGAAACA  TAAAAATCTT  CTGTACTTCA  CTCCAAGGTA  2580
      CATTTGCTTA  CTGACAGCAT  TTTTGTTAAA  ACTGTTATTC  TTGAAAAAAA  AAAAAAAAAA  2640
      AA
```

Seq ID NO: 90 Protein sequence
Protein Accession #: AAC39582.1

```
      1           11          21          31          41          51
      |           |           |           |           |           |
70    KEAASGKSDR  RTPESSLHEA  LDQCMTALDL  PLTNQFSEAL  SYLKPRTKES  MYHSLTYATI   60
      LEMQAMMTFD  PQDILLAGNM  MKEAQMLCQR  HRRKSSVTDS  FSSLVNRPTL  GQFTEEEIHA  120
      EVCYAECLLQ  RAALTFLQDE  NMVSPIKGGI  KVRNSYQTYK  ELDSLVQSSQ  YCKGENHPHF  180
      EGGVKLGVGA  FNLTLSMLPT  RILRLLEFVG  FSGNKDYGLL  QLEEGASGHS  FRSVLCVMLL  240
75    LCYHTFLTFV  LGTGNVNIEE  AEKLLKPYLN  RYPKGAIFLP  FAGRIEVIKG  NIDAAIRRFE  300
      ECEAQQHWK   QFHHMCYWEL  MWCFTYKGQW  KMSYFYADLL  SKENCWSKAT  YIYMKAAYLS  360
      MFGKEDHKPF  GDDEVELFRA  VPGLKLKIAG  KSLPTEKFAI  RKSRRYFSSN  PISLPVPALE  420
      MMYIWNGYAV  IGKQPKLTDG  ILEIITKAEE  MLEKGPENEY  SVDDECLVKL  LKGLCLKYLG  480
      RVQEAEENFR  SISANEKKIK  YDHYLIPNAL  LELALLLMEQ  DRNEEAIKLL  ESAKQNYKNY  540
      SMESRTHFRI  QAATLQAKSS  LENSSRSMVS  SVSL
```

Seq ID NO: 91 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 31..906

```
      1         11        21        31        41        51
      |         |         |         |         |         |
      CGGGTCGACC CACGCGTCCG GGGAGAAAGG ATGGCCGGCC TGGCGGCGCG GTTGGTCCTG    60
      CTAGCTGGGG CAGCGGCGCT GGCGAGCGGC TCCCAGGGCG ACCGTGAGCC GGTGTACCGC   120
      GACTGCGTAC TGCAGTGCGA AGAGCAGAAC TGCTCTGGGG GCGCTCTGAA TCACTTCCGC   180
      TCCCGCCAGC CAATCTACAT GAGTCTAGCA GGCTGGACCT GTCGGGACGA CTGTAAGTAT   240
      GAGTGTATGT GGGTCACCGT TGGGCTCTAC CTCCAGGAAG GTCACAAAGT GCCTCAGTTC   300
      CATGGCAAGT GGCCCTTCTC CCGGTTCCTG TTCTTTCAAG AGCCGGCATC GGCCGTGGCC   360
      TCGTTTCTCA ATGGCCTGGC CAGCCTGGTG ATGCTCTGCC GCTACCGCAC CTTCGTGCCA   420
      GCCTCCTCCC CCATGTACCA CACCTGTGTG GCCTTCGCCT GGGTGTCCCT CAATGCATGG   480
      TTCTGGTCCA CAGTTTCCA CACCAGGGAC ACTGACCTCA CAGAGAAAT GGACTACTTC   540
      TGTGCCTCCA CTGTCATCCT ACACTCAATC TACCTGTGCT GCGTCAGCCT CATCCGCTTC   600
      GACTATGGCT ACAACCTGGT GGCCAACGTG GCTATTGGCC TGGTCAACGT GGTGTGGTGG   660
      CTGGCCTGGT GCCTGTGGAA CCAGCGGCGG CTGCCTCACG TGCGCAAGTG CGTGGTGGTG   720
      GTCTTGCTGC TGCAGGGGCT GTCCCTGCTC GAGCTGCTTG ACTTCCCACC GCTCTTCTGG   780
      GTCCTGGATG CCCATGCCAT CTGGCACATC AGCACCATCC CTGTCCACGT CCTCTTTTTC   840
      AGCTTTCTGG AAGATGACAG CCTGTACCTG CTGAAGGAAT CAGAGGACAA GTTCAAGCTG   900
      GACTGAAGAC CTTGGAGCGA GTCTGCCCCA GTGGGGATCC TGCCCCCGCC CTGCTGGCCT   960
      CCCTTCTCCC CTCAACCCTT GAGATGATTT TCTCTTTTCA ACTTCTTGAA CTTGGACATG  1020
      AAGGATGTGG GCCCAGAATC ATGTGGCCAG CCCACCCCCT GTTGGCCCTC ACCAGCCTTG  1080
      GAGTCTGTTC TAGGGAAGGC CTCCCAGCAT CTGGGACTCG AGAGTGGGCA GCCCCTCTAC  1140
      CTCCTGGAGC TGAACTGGGG TGGAACTGAG TGTGCTCTTA GCTCTACCGG GAGGACAGCT  1200
      GCCTGTTTCC TCCCCATCAG CCTCCTCCCC ACATCCCCAG CTGCCTGGCT GGGTCCTGAA  1260
      GCCCTCTGTC TACCTGGGAG ACCAGGGACC ACAGGCCTTA GGGTCCCCTT  1320
      CTGTTACCAC CCCCCACCCT CCTCCAGGAC ACCACTAGGT GGTGCTGGAT GCTTGTTCTT  1380
      TGGCCAGCCA AGGTTCACGG CGATTCTCCC CATGGGATCT TGAGGGACCA AGCTGCTGGG  1440
      ATTGGGAAGG AGTTTCACCC TGACCRTTGC CCTAGCCAGG TTCCCAGGAG GCCTCACCAT  1500
      ACTCCCTTTC AGGGCCAGGG CTCCAGCAAG CCCAGGGCAA GGATCCTGTG CTGCTGTCTG  1560
      GTTGAGAGCC TGCCACCGTG TGTCGGGAGT GTGGGCCAGG CTGAGTGCAT AGGTGACAGG  1620
      GCCGTGAGCA TGGGCCTGGG TGTGTGTGAG CTCAGGCACT AGGTGCGCAG TGTGGAGACG  1680
      GGTGTTGTCG GGGAAGAGGT GTGGCTTCAA AGTGTTGTGT GTGCAGGGGG TKGGTGTGTT  1740
      AAGCGTGGGT TAGGGGAACG TGTGTGCGCG TGCTGGTGGG CATGTGAGAT GAGTGACTGC  1800
      CGGTGAATGT GTCCACAGTT GAGAGGTTGG AGCAGGATGA GGGAATCCTG TCACCATCAA  1860
      TAATCACTTG TGGAGCGCCA CTTGGCCCAA GACGCCACCT GGGCGGACAG CAGGAGCTCT  1920
      CCATGGCCAG GCTGCCTGTG TGCATGTTCC CTGTCTGGTG CCCCTTTGCC CGCCTCCTGC  1980
      AAACCTCACA GGGTCCCCAC ACAACAGTGC CCTCCAGAAG CAGCCCCTCG GAGGCAGAGG  2040
      AAGGAAAATG GGGATGGCTG GGGCTCTCTC CATCCTCCTT TTCTCCTTGC CTTCGCATGG  2100
      CTGGCCTTCC CCTCCAAAAC CTCCATTCCC CTGCTGCCAG CCCCTTTGCC ATAGCCTGAT  2160
      TTTGGGGAGG AGGAAGGGGC GATTTGAGGG AGAAGGGGAG AAAGCTTATG GCTGGGTCTG  2220
      GTTTCTTCCC TTCCCAGAGG GTCTTACTGT TCCAGGGTGG CCCCAGGGCA GGCAGGGGCC  2280
      ACACTATGCC TGCGCCCTGG TAAAGGTGAC CCCTGCCATT TACCAGCAGC CCTGGCATGT  2340
      TCCTGCCCCA CAGGAATAGA ATGGAGGGAG CTCCAGAAAC TTTCCATCCC AAAGGCAGTC  2400
      TCCGTGGTTG AAGCAGACTG GATTTTTGCT CTGCCCCTGA CCCCTTGTCC CTCTTTGAGG  2460
      GAGGGGAGCT ATGCTAGGAC TCCAACCTCA GGGACTCGGG TGGCCTGCGC TAGCTTCTTT  2520
      TGATACTGAA AACTTTTAAG GTGGGAGGGT GGCAAGGGAT GTGCTTAATA AATCAATTCC  2580
      AAGCCTCAAA AAAAAAAAAA AAAAAAAAA AAAAA
```

Seq ID NO: 92 Protein sequence
Protein Accession #: Eos sequence

```
      1         11        21         31        41        51
      |         |         |          |         |         |
      MAGLAARLVL LAGAAALASG SQGDREPVYR DCVLQCEEQN CSGGALNHFR SRQPIYMSLA    60
      GWTCRDDCKY ECMWVTVGLY LQEGHKVPQF HGKWPFSRFL FFQEPASAVA SFLNGLASLV   120
      MLCRYRTFVP ASSPMYHTCV APAWVSLNAW FWSTVPHTRD TDLTEKMDYF CASTVILHSI   180
      YLCCVRTVGL QHPAVVSAFR ALLLLMLTVH VSYLSLIRFD YGYNLVANVA IGLVNVVWWL   240
      AWCLWNQRRL PHVRKCVVVV LLLQGLSLLE LLDFPPLFWV LDAHAIWHIS TIPVHVLFFS   300
      FLEDDSLYLL KESEDKFKLD
```

Seq ID NO: 93 DNA sequence
Nucleic Acid Accession #: NM_033419.1
Coding sequence: 18..980

```
      1         11        21        31        41        51
      |         |         |         |         |         |
      CGAGCCAGGG AGAAAGGATG GCCGGCCTGG CGGCGCGGTT GGTCCTGCTA GCTGGGGCAG    60
      CGGCGCTGGC GAGCGGCTCC CAGGGCGACC GTGAGCCGGT GTACCGCGAC TGCGTACTGC   120
      AGTGCGAAGA GCAGAACTGC TCTGGGGGCG CTCTGAATCA CTTCCGCTCC CGCCAGCCAA   180
      TCTACATGAG TCTAGCAGGC TGGACCTGTC GGGACGACTG TAAGTATGAG TGTATGTGGG   240
      TCACCGTTGG GCTCTACCTC CAGGAAGGTC ACAAAGTGCC TCAGTTCCAT GGCAAGTGGC   300
      CCTTCTCCCG GTTCCTGTTC TTTCAAGAGC CGGCATCGGC CGTGGCCTCG TTTCTCAATG   360
      GCCTGGCCAG CCTGGTGATG CTCTGCCGCT ACCGCACCTT CGTGCCAGCC TCCTCCCCCA   420
      TGTACCACAC CTGTGTGGCC TTCGCCTGGG TGTCCCTCAA TGCATGGTTC TGGTCCACAG   480
      TTTTCCACAC CAGGGACACT GACCTCACAG AGAAAATGGA CTACTTCTGT GCCTCCACTG   540
      TCATCCTACA CTCAATCTAC CTGTGCTGCG TCAGGACCGT GGGGCTGCAG CACCCAGCTG   600
      TGGTCAGTGC CTTCCGGGCT CTCCTGCTGC TCATGCTGAC CGTGCACGTC TCCTACCTGA   660
      GCCTCATCCG CTTCGACTAT GGCTACAACC TGGTGGCCAA CGTGGCTATT GGCCTGGTCA   720
      ACGTGGTGTG GTGGCTGGCC TGGTGCCTGT GGAACCAGCG GCGGCTGCCT CACGTGCGCA   780
      AGTGCGTGGT GGTGGTCTTG CTGCTGCAGG GGCTGTCCCT GCTCGAGCTG CTTGACTTCC   840
      CACCGCTCTT CTGGGTCCTG GATGCCCATG CCATCTGGCA CATCAGCACC ATCCCTGTCC   900
      ACGTCCTCTT TTTCAGCTTT CTGGAAGATG ACAGCCTGTA CCTGCTGAAG GAATCAGAGG   960
      ACAAGTTCAA GCTGGACTGA AGACCTTGGA GCGAGTCTGC CCCAGTGGGG ATCCTGCCCC  1020
      CGCCCTGCTG GCCTCCCTTC TCCCCTCAAC CCTTGAGATG ATTTTCTCTT TTCAACTTCT  1080
      TGAACTTGGA CATGAAGGAT GTGGGCCCAG AATCATGTGG CCAGCCCACC CCCTGTTGGC  1140
```

```
CCTCACCAGC CTTGGAGTCT GTTCTAGGGA AGGCCTCCCA GCATCTGGGA CTCGAGAGTG  1200
GGCAGCCCCT CTACCTCCTG GAGCTGAACT GGGGTGGAAC TGAGTGTGCT CTTAGCTCTA  1260
CCGGGAGGAC AGCTGCCTGT TTCCTCCCCA TCAGCCTCCT CCCCACATCC CCAGCTGCCT  1320
GGCTGGGTCC TGAAGCCCTC TGTCTACCTG GGAGACCAGG GACCACAGGC CTTAGGGATA  1380
CAGGGGGTCC CCTTCTGTTA CCACCCCCCA CCCTCCTCCA GGACACCACT AGGTGGTGCT  1440
GGATGCTTGT TCTTTGGCCA GCCAAGGTTC ACGGCGATTC TCCCCATGGG ATCTTGAGGG  1500
ACCAAGCTGC TGGGATTGGG AAGGAGTTTC ACCCTGACCA TTGCCCTAGC CAGGTTCCCA  1560
GGAGGCCTCA CCATACTCCC TTTCAGGGGC AGGGCTCCAG CAAGCCCAGG GCAAGGATCC  1620
TGTGCTGCTG TCTGGTTGAG AGCCTGCCAC CGTGTGTCGG GAGTGTGGGC CAGGCTGAGT  1680
GCATAGGTGA CAGGGCCGTG AGCATGGGCC TGGGTGTGTG TGAGCTCAGG CCTAGGTGCG  1740
CAGTGTGGAG ACGGGTGTTG TCGGGGAAGA GGTGTGGCTT CAAAGTGTGT GTGTGCAGGG  1800
GGTGGGTGTG TTAGCGTGTGGG TAGGGGAAC GTGTGTGCGC GTGCTGGTGG GCATGTGAGA  1860
TGAGTGACTG CCGGTGAATG TGTCCACAGT TGAGAGGTTG GAGCAGGATG AGGGAATCCT  1920
GTCACCATCA ATAATCACTT GTGGAGCGCC AGCTCTGCCC AAGGCGCCAC CTGGGCGGAC  1980
AGCCAGGAGC TCTCAAACCTGC CAGGCTGCCT GTGTGCATGT TCCCTGTCTG GTGCCCCTTT  2040
GCCCGCCTCC TGCAAACCTC ACAGGGTCCC CACACAACAG TGCCCTCCAG AAGCAGCCCC  2100
TCGGAGGCAG AGGAAGGAAA ATGGGGATGG CTGGGGCTCT CTCCATCCTC CTTTTCTCCT  2160
TGCCTTCGCA TGGCTGGCCT TCCCCTCCAA AACCTCCATT CCCCTGCTGC CAGCCCCTTT  2220
GCCCATAGCCT GATTTGGGG AGGAGGAAGG GGCGATTGA GGGAGAAGGG GAGAAAGCTT  2280
ATGGCTGGGT CTGGTTTCTT CCCTTCCCAG AGGGTCTTAC TGTTCCAGGG TGGCCCCAGG  2340
GCAGGCAGGG GCCACACTAT GCCTGCGCCC TGGTAAAGGT GACCCCTGCC ATTTACCAGC  2400
AGCCCTGGCA TGTTCCTGCC CCACAGGAAT AGAATGGAGG GAGCTCCAGA AACTTTCCAT  2460
CCCAAAGGCA GTCTCCGTGG TTGAAGCAGA CTGGATTTTT GCTCTGCCCC TGACCCCTTG  2520
TCCCTCTTTG AGGGAGGGGA GCTATGCTAG GACTCCAACC TCAGGGACTC GGGTGGCCTG  2580
CGCTAGCTTC TTTTGATACT GAAAACTTTT AAGGTGGGAG GGTGGCAAGG GATGTGCTTA  2640
ATAAATCAAT TCCAAGCCTC AAAAAAAAAA AAAAAAAAA AA
```

Seq ID NO: 94 Protein sequence
Protein Accession #: NP_219487.1

```
1         11        21        31        41        51
|         |         |         |         |         |
MKDVGPESCG QPTPCWPSPA LESVLGKASQ HLGLESGQPL YLLELNWGGT ECALSSTGRT   60
AACFLPISLL PTSPAAWLGP EALCLPGRPG TTGLRDTGGP LLLPPPTLLQ DTTRWCWMLV  120
LWPAKVHGDS PHGILRDQAA GIGKEFHPDH CPSQVPRRPH HTPFQGQGSS KPRARILCCC  180
LVESLPPCVG SVGQAECIGD RAVSMGLGVC ELRPRCAVWR RVLSGKRCGF KVCVCRGWVC
```

Seq ID NO: 95 DNA sequence
Nucleic Acid Accession #: XM_090469
Coding sequence:

```
1         11        21        31        41        51
|         |         |         |         |         |
ATGGGGTTTG GAGACCAGGG AACGGTGGAA GGGAGCCTAG GAACGTCGAA AGATAGCTCC   60
TGCCGGCGTGC TGGACAGGAC GCAGGTGAAC AAGCCCTGCC TGCTGGAGGC TGCCAAGCAC  120
AGCCCACACC CCTTGCATGG ACAGTCGGTC AACAGCCACC CATCAGGTGC TCATCAGAAA  180
CCACCTGAAG TGAAAATGTT TGGAGCCAGT CAAGGTTTGC TGACAATGGA AACAAACCAG  240
TCCCTGGCAC AAGGCACAGG CTGCTCAGTG TGCTGGGAGG TGAACGGCCC AGGGTTGGAA  300
ATGTCTCTTC CTGGAGTGCT CAGTGCTGAT GCTGGCCAGG TGGAGCACAG AAGACAAATG  360
AACTCCGCAG ATACTGACTG GGGAACTGAG GAGGGGACCAG GAAGTCTAGA TTTCGCCGTG  420
GCTGCCCACC AGGAAGACAC TTTCTTTTTG AAAGACATCA AGCACACAAG TACGTTCAGG  480
CAATCAGTTC AGCAGCAGAA TTGCATTTAC AGCCCCAGAG AAAAACCCTG TGGGAATGTC  540
AGGGCACCTT GCGCGCCACC AAGGCGAGAG CCCTCAGCAG GCCCTGGCGG  600
CTCAGCCGCA GCCCGGCGCC GTCCCCACGT ACACCCATGG CGCCCTTCCC GACCTCTGAC  660
CGCGAGCTAG ACGCGGCCAGG CCCGCCCCCC GGGCTGCGGA GCTCGGCAGC AGCCCCTCAC  720
TGCCTGCCCG CGGCTCCAGA AGCCCAGAAA CTCTTCTTCC TGCTGCCCCT TTATCCAGAT  780
GGCAGCCCAC CACCTAAGGA CATCTTGCAG ACGCTCCAGC ACAAGCCACC TGGGAAGAGA  840
CTGCATCAAG AATCACCAGG TTCATTTCCA ATAGGTTTAT GTAACAGAAC AGCTGAACCA  900
TGTATCAAAG GGGCTCTGGG TGTGGCTGCT GAAGCAGCAT TTCACTTACA GTTTTCCAGT  960
GAATGA
```

Seq ID NO: 96 Protein sequence
Protein Accession #: XP_090469

```
1         11        21        31        41        51
|         |         |         |         |         |
MGFGDQGTVE GSLGTSKDSS CRVLDRTQVN KPCLLEAAKH SPHPLHGQSV NSHPSGAHQK   60
PPEVKMFGAS QGLLTMETNQ SLAQGTGCSV CWEVNGPGLE MSLPGVLSAD AGQVEHRRQM  120
NSADTDWGTE EGPGSLDFAV AAHQEDTFFL KDIKHTSFR QSVQQQNCIY SPREKPCGNV  180
RAPCAPPRRE APLALSRPWR LSRSPAPSPR TPMAPFPTSD RELDAPGPPP GLRSSAAAPH  240
CLPAAPEAQK LFFLLPLYPD GSPPPKDILQ TLQHKAPGKR LHQESPGSFP IGLCNRTAEP  300
CIKGALGVAA EAAFHLQFSS E
```

Seq ID NO: 97 DNA sequence
Nucleic Acid Accession #: NM_003474.2
Coding sequence: 37..3036

```
1         11        21        31        41        51
|         |         |         |         |         |
CACTAACGCT CTTCCTAGTC CCCGGGCCAA CTCGGACAGT TTGCTCATTT ATTGCAACGG   60
TCAAGGCTGG CTTGTGCCAG AACGGCGCGC GCGCGGACGCA CGCACACACA CGGGGGGAAA  120
CTTTTTTAAA AATGAAAGGC TAGAAGAGCT CAGCGGCGGC GCGGCGCGTG CGCGAGGGCT  180
CGGGAGCTGA CTCGCCGAGG CAGGAAATCC CTCCGGTCGC GACGGCCCGG CCCGCTCGGC  240
GCCCGCGTGG GATGGTGCAG CGCTCGCCGC CGGGCCCGAG AGCTGCTGCA CTGAAGGCCG  300
```

```
        GCGACGATGG CAGCGCBCCC GCTGCCCGTG TCCCCCGCCC GCGCCCTCCT GCTCGCCCTG   360
        GCCGGTGCTC TGCTCGCGCC CTGCGAGGCC CGAGGGGTGA GCTTATGGAA CGAAGGAAGA   420
        GCTGATGAAG TTGTCAGTGC CTCTGTTCGG AGTGGGGACC TCTGGATCCC AGTGAAGAGC   480
        TTCGACTCCA AGAATCATCC AGAAGTGCTG AATATTCGAC TACAACGGGA AAGCAAAGAA   540
   5    CTGATCATAA ATCTGGAAAG AAATGAAGGT CTCATTGCCA GCAGTTTCAC GGAAACCCAC   600
        TATCTGCAAG ACGGTACTGA TGTCTCCCTC GCTCGAAATT ACACGGTAAT TCTGGGTCAC   660
        TGTTACTACC ATGGACATGT ACGGGGGATAT TCTGATTCAG CAGTCAGTCT CAGCACGTGT   720
        TCTGGTCTCA GGGGACTTAT TGTGTTTGAA AATGAAAGCT ATGTCTTAGA ACCAATGAAA   780
        AGTGCAACCA ACAGATACAA ACTCTTCCCA GCGAAGAAGC TGAAAAGCGT CCGGGGGATCA   840
   10   TGTGGATCAC ATCACAACAC ACCAAACCTC GCTGCAAAGA ATGTGTTTCC ACCACCCTCT   900
        CAGACATGGG CAAGAAGGCA TAAAAGAGAG ACCCTCAAGG CAACTAAGTA TGTGGAGCTG   960
        GTGATCGTGG CAGACAACCG AGAGTTTCAG AGGCAAGGAA AAGATCTGGA AAAAGTTAAG   1020
        CAGCGATTAA TAGAGATTGC TAATCACGTT GACAAGTTTT ACAGACCACT GAACATTCGG   1080
        ATCGTGTTGG TAGGCGTGGA AGTGTGGAAT GACATGGACA AATGCTCTGT AAGTCAGGAC   1140
   15   CCATTCACCA GCCTCCATGA ATTTCTGGAC TGGAGGAAGA TGAAGCTTCT ACCTCGCAAA   1200
        TCCCATGACA ATGCGCAGCT TGTCAGTGGG GTTTATTTCC AAGGGACCAC CATCGGCATG   1260
        GCCCCAATCA TGAGCATGTG CACGGCAGAC CAGTCTGGGG GAATTGTCAT GGACCATTCA   1320
        GACAATCCCC TTGGTGCAGC CGTGACCCTG GCACATGAGC TGGGCCACAA TTTCGGGATG   1380
        AATCATGACA CACTGGACAG GGGCTGTAGC TGTCAAATGG CGGTTGAGAA AGGAGGCTGC   1440
   20   ATCATGAACG CTTCCACCGG GTACCCATTT CCCATGGTGT TCAGCAGTTG CAGCAGGAAG   1500
        GACTTGGAGA CCAGCCTGGA GAAAGGAATG GGGGTGTGCC TGTTTAACCT GCCAGGAAGTC   1560
        AGGGAGTCTT TCGGGGGCCA GAAGTGTGGG AACAGATTTG TGGAAGAAGG AGAGGAGTGT   1620
        GACTGTGGGG AGCCAGAGGA ATGTATGAAT CGCTGCTGCA ATGCCACCAC CTGTACCCTG   1680
        AAGCCGGACG CTGTGTGCGC ACATGGGCTG TGCTGTGAAG ACTGCCAGCT GAAGCCTGCA   1740
   25   GGAACAGCGT GCAGGGACTC CAGCAACTCC TGTGACCTCC CAGAGTTCTG CACAGGGGCC   1800
        AGCCCTCACT GCCCAGCCAA CGTGTACCTG CACGATGGGC ACTCATGTCA GGATGTGGAC   1860
        GGCTACTGCT ACAATGGCAT CTGCCAGACT CACGAGCAGC AGTGTGTGCA ACTCTGGGGA   1920
        CCAGGTGCTA AACCTGCCCC TGGGATCTGC TTTGAGAGAG TCAATTCTGC AGGTGATCCT   1980
        TATGGCAACT GTGGCAAAGT CTCGAAGAGT TCCTTTGCCA AATGCGAGAT GAGAGATGCT   2040
   30   AAATGTGGAA AAATCCAGTG TCAAGGAGGT GCCAGCCGGC CAGTCATTGG TACCAATGCC   2100
        GTTTCCATAG AAACAAACAT CCCCCTGCAG CAAGGAGGCC GGATTCTGTG CCGGGGGACC   2160
        CACGTGTACT TGGGCGATGA CATGCCGGAC CCAGGGCTTG TGCTTGCAGG CACAAAGTGT   2220
        GCAGATGGAA AAATCTGCCT GAATCGTCAA TGTCAAAATA TTAGTGTCTT TGGGGTTCAC   2280
        GAGTGTGCAA TGCCAGTGCCA CGGCAGAGGG GTGTGCAACA ACAGGAAGAA CTGCCACTGC   2340
   35   GAGGCCCACT GGGCACCTCC CTTCTGTGAC AAGTTTGGCT TTGGAGGAAG CACAGACAGC   2400
        GGCCCCATCC GGCAAGCAGA TAACCAAGGT TTAACCATAG GAATTCTGGT GACCATCCTG   2460
        TGTCTTCTTG CTGCCGGGATT TGTGGTTTAT CTCAAAAGGA AGACCTTGAT ACGACTGCTG   2520
        TTTACAAATA AGAAGACCAC CATTGAAAAA CTAAGGTGTG TGCGCCCTTC CCGGCCACCC   2580
        CGTGGCTTCC AACCCTGTCA GGCTCACCTC GGCCACCTTG GAAAAGGCCT GATGAGGAAG   2640
   40   CCGCCAGATT CCTACCCACC GAAGGACAAT CCCAGGAGAT TGCTGCAGTG TCAGAATGTT   2700
        GACATCAGCA GACCCCTCAA CGGCCTGAAT GTCCCTCAGC CCCAGTCAAC TCAGCGAGTG   2760
        CTTCCTCCCC TCCACCGGGC CCCACGTGCA CCTAGCGTCC CTGCCAGACC CCTGCCAGCC   2820
        AAGCCTGCAC TTAGGCAGGC CCAGGGGACC TGTAAGCCAA ACCCCCCTCA GAAGCCTCTG   2880
        CCTGCAGATC CTCTGGCCAG AACAACTCGG CTCACTCATG CCTTGGCCAG GACCCCAGGA   2940
   45   CAATGGGAGA CTGGGCTCCG CCTGGCACCC CTCAGACCTG CTCCACAATA TCCACACCAA   3000
        GTGCCAGAT CCACCCACAC CGCCCTATATT AAGTGAGAGG CCGACACCTT TTTTCAACAG   3060
        TGAAGACAGA AGTTTGCACT ATCTTTCAGC TCCAGTTGGA GTTTTTTGTA CCAACTTTTA   3120
        GGATTTTTTT TAATGTTTAA AACATCCATTA CTATAAGAAC TTTGAGCTAC TGCCGTCAGT   3180
        GCTGTGCTGT GCTATGGTGC TCTGTCTACT TGCACAGGTA CTTGTAAATT ATTAATTTAT   3240
   50   GCAGAATGTT GATTACAGTG CAGTGCGCTG TAGTAGGCAT TTTTACCATC ACTGAGTTTT   3300
        CCATGGCAGG AAGGCTTGTT GTGCTTTTAG TATTTTAGTG AACTTGAAAT ATCCTGCTTG   3360
        ATGGGATTCT GGACAGGATG TGTTTGCTTT CTGATCAAGG CCTTATTGGA AAGCAGTCCC   3420
        CCAACTACAG CCAGCTGTGC TTATGGTACC AGATGCAGCT CAAGAGATCC CAAGTAGAAT   3480
   55   CTCAGTTGAT TTTCTGGATT CCCCATCTCA GGCCAGAGCC AAGGGGCTTC AGGTCCAGGC   3540
        TGTGTTTGGC TTTCAGGGAG GCCCTGTGCC CCTTGACAAC TGGCAGGCAG GCTCCCAGGG   3600
        ACACCTGGGA GAAATCTGGC TTCTGGCCAG GAAGCTTTGG TGAGAACCTG GGTTGCAGAC   3660
        AGGAATCTTA AGGTGTAGCC ACACCAGGAT AGAGACTGGA ACACTAGACA AGCCAGAACT   3720
        TGACCCTGAG CTGACCAGCC GTGAGCATGT TTGGAAGGGG TCTGTAGTGT CACTCAAGGC   3780
   60   GGTGCTTGAT AGAAATGCCA AGCACTTCTT TTTCTCGCTG TCCTTTCTAG GCACTGCCCA   3840
        CCAGTAGGTT ATTTAGCTTG GGAAAGGTGG TGTTTCTGTA AGAAACCTAC TGCCCAGGCA   3900
        CTGCAAACCG CCACCTCCCT ATACTGCTTG GAGCTGAGCA AATCACCACA AACTGTAATA   3960
        CAATGATCCT GTATTCAGAC AGATGAGGAC TTTCCATGGC ACCACAACTA TTTTCAGATG   4020
        TGAACCATTA ACCAGATCTA GTCAATCAAG TCTGTTTACT GCAAGGTTCA ACTTATTAAC   4080
   65   AATTAGGCAG ACTCTTTATG CTTGCAAAAA CTACAACCAA TGGAATGTGA TGTTCATGGG   4140
        TATAGTTCAT GTCTGCTATC ATTATTCGTA GATATTGGAC AAAGAACCTT CCTCATGGGG   4200
        CATCCTCTTT TTCCAACTTG GCTGCAGGAA TCTTTAAAAG ATGCTTTTAA CAGAGTCTGA   4260
        ACCTATTTCT TAAACACTTG CAACCTACCT GTTGAGCATC ACAGAATGTG ATAAGGAAAT   4320
        CAACTTGCTT ATCAACTTCC TAAATATTAT GAGATGTGTC TTGGGCAGCA TCCCCTGAAA   4380
   70   CTCTTCACTC TTCAAATGCC TGACTAGGGA GCCATGTTTC ACAAGGTCTT TAAAGTGACT   4440
        AATGGCATGA GAAATACAAA AATACTCAGA TAAGGTAAAA TGCCATGATG CCTCTGTCTT   4500
        CTGGACTGGT TTTCACATTA GAAGACAATT GACAACAGTT ACATAATTCA CTCTGAGTGT   4560
        TTTATGAGAA AGCCTTCTTT TGGGGTCAAC AGTTTTCCTA TGCTTTGAAA CAGAAAAATA   4620
        TGTACCAAGA ATCTTGGTTT GCCTCCAGA AAACAAAACT GCATTTCACT TTCCCGGTGT   4680
   75   TCCCCACTGT ATCTAGGCAA CATAGTATTC ATGACTATGG ATAAACTAAA CACGTGACAC   4740
        AAACACACAC AAAAGGGAAC CCAGCTCTAA TACATTCCAA CTCGTATAGC ATGCATCTGT   4800
        TTATTCTATA GTTATTAAGT TCTTTAAAAT GTAAAGCCAT GCTGGAAAAT AATACTGCTG   4860
        AGATACATAC AGAATTACTG TAACTGATTA CACTTGGTAA TTGTACTAAA GCCAAACATA   4920
        TATATACTAT TAAAAAGGTT TACAGAATTT TATGGTGCAT TACGTGGGCA TTGTCTTTTT   4980
        AGATGCCCAA ATCCTTAGAT CTGGCATGTT AGCCCTTCCT CCAATTATAA GAGGGATATGA   5040
   80   ACCAAAAAAA AAAAAAAAAA AA
```

Seq ID NO: 98 Protein sequence
Protein Accession #: NP_003465

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        MAARRPLPVSP ARALLLALAG ALLAPCEARG VSLWNEGRAD EVVSASVRSG DLWIPVKSFD   60
        SKNHPEVLNI RLQRESKELI INLERNEGLI ASSFTETHYL QDGTDVSLAR NYTVILGHCY   120
        YHGHVRGYSD SAVSLSTCSG LRGLIVPENE SYVLEPMKSA TNRYKLPPAK KLKSVRGSCG   180
        SHHNTPNLAA KNVFPPPSQT WARRHKRETL KATKYVELVI VADNREFQRQ GKDLEKVKQR   240
        LIEIANHVDK FYRPLNIRIV LVGVEVWNDM DXCSVSQDPF TSLHEFLDHR KMKLLPRKSH   300
        DNAQLVSGVY FQGTTIGMAP IMSMCTADQS GGIVMDHSDN PLGAAVTLAH ELGHNFGMNH   360
        DTLDRGCSCQ MAVEKGGCIM NASTGYPFPM VFSSCSRKDL RTSLEKGMGV CLFNLPEVRE   420
        SFGGQKCGNR FVEEGEECDC GEPEECMNRC CNATTCTLKP DAVCAHGLCC EDCQLKPAGT   480
        ACRDSSNSCD LPEPCTGASP HCPANVYLHD GHSCQDVDGY CYNGICQTHE QQCVTLWGPG   540
        AKPAPGICFE RVNSAGDPYG NCGKVSKSSF AKCEMRDAKC GKIQCQGGAS RPVIGTNAVS   600
        IETNIPLQQG GRILCRGTHV YLGDDMPDPG LVLAGTKCAD GKICLNRQCQ NISVFGVREC   660
        AMQCHGRGVC NNRKNCHCEA HWAPPFCDKF GFGGSTDSGP IRQADNQGLT IGILVTILCL   720
        LAAGPVVYLK RKTLIRLLFT NKKTTIEKLR CVRPSRPPRG FQPCQAHLGH LGKGLMRKPP   780
        DSYPPKDNPR RLLQCQNVDI SRPLNGLNVP QPQSTQRVLP PLHRAPRAPS VPARPLPAKP   840
        ALRQAQGTCK PNPPQKPLPA DPLARTTRLT HALARTPGQW ETGLRLAPLR PAPQYPHQVP   900
        RSTHTAYIK
```

Seq ID NO: 99 DNA sequence
Nucleic Acid Accession #: NM_003714
Coding sequence: 135..1043

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        GAGGAGGAGG GAAAAGGCGA GCAAAAAGGA AGAGTGGGAG GAGGAGGGGA AGCGGCGAAG   60
        GAGGAAGAGG AGGAGGAGGC AGAGGGGCGC ACAAAGGATC CAGGTCTCCC GACGGGAGGT   120
        TAATACCAAG AACCATGTGT GCCGAGCGGC TGGGCCAGTT CATGACCCTG GCTTTGGTGT   180
        TGGCCACCTT TGACCCGGCG CGGGGGACCG ACGCCACCAA CCCACCCGAG GGTCCCCAAG   240
        ACAGGAGCTC CCAGCAGAAA GGCCGCCTGT CCCTGCAGAA TACAGCGGAG ATCCAGCACT   300
        GTTTGGTCAA CGCTGGCGAT GTGGGGTGTG GCGTGTTTGA ATGTTTCGAG AACAACTCTT   360
        GTGAGATTCG GGGCTTACAT GGGATTTGCA TGACTTTTCT GCACAACGCT GGAAAATTTG   420
        ATGCCCAGGG CAAGTCATTC ATCAAAGACG CCTTGAAATG TAAGGCCCAC GCTCTGCGGC   480
        ACAGGTTCGG CTGCATAAGC CGGAAGTGCC CGGCCATCAG GGAAATGGTG TCCCAGTTGC   540
        AGCGGGAATG CTACCTCAAG CACGACCTGT GCGCGGCTGC CCAGGAGAAC ACCGGGGTGA   600
        TAGTGGAGAT GATCCATTTC AAGGACTTGC TGCTGCACAA ACCCTACGTG GACCTCGTGA   660
        ACTTGCTGCT GACCCTGTGG GAGGAGGTGA AGGAGGCCAT CACCCACAGC GTGCAGGTTC   720
        AGTGTGAGCA GAACTGGGGA AGCCTGTGCT CCATCTTGAG CTTCTGCACC TCGGCCATCC   780
        AGAAGCCTCC CACGGCGCCC CCCGAGCGGC AGCCCCAGGT GGACAGAACC AAGCTCTCCA   840
        GGGCCCACCA CGGGGAAGCA GGACATCACC TCCCAGAGCC CAGCAGTAGG GAGACTGGCC   900
        GAGGTGCCAA GGGTGAGCGA GGTAGCAAGA GCCACCCAAA CGCCCATGCC CGAGGCAGAG   960
        TCGGGGGCCT TGGGGCTCAG GGACCTTCCG GAAGCAGCGA GTGGGAAGAC GAACAGTCTG   1020
        AGTATTCTGA TATCGGGAGG TGAAATGAAA GGCCTGGCCA CGAAATCTTT CCTCCACGCC   1080
        GTCCATTTTC TTATCTATGG ACATTCCAAA ACATTTACCA TTAGAGAGGG GGGATGTCAC   1140
        ACGCAGGATT CTGTGGGGAC TGTGGACTTC ATCGAGGTGT GTGTTCGCGG AACGGACAGG   1200
        TGAGATGGAG ACCCCTGGGG CCGTGGGGTC TCAGGGGTGC CTGGTGAATT CTGCACTTAC   1260
        ACGTACTCAA GGGAGCGCGC CCGCGTTATC CTCGTACCTT TGTCTTCTTT CCATCTGTGG   1320
        AGTCAGTGGG TGTCGGCCGC TCTGTTGTGG GGGAGGTGAA CCAGGGAGGG GCAGGGCAAG   1380
        GCAGGGCCCC CAGAGCTGGG CCACACAGTG GGTGCTGGGC CTCGCCCCGA AGCTTCTGGT   1440
        GCAGCAGCCT CTGGTGCTGT CTCCGCGGAA GTCAGGGCGG CTGGATTCCA GGACAGGAGT   1500
        GAATGTAAAA ATAAATATCG CTTAGAATGC AGGAGAAGGG TGGAGAGGAG GCAGGGGCCG   1560
        AGGGGGTGCT TGGTGCCAAA CTGAAATTCA GTTTCTTGTG TGGGGCCTTG CGGTTCAGAG   1620
        CTCCTTGGCGA GGGTGGAGGG AGGAGTGTCA TTTCTATGTG TAATTTCTGA GCCATTGTAC   1680
        TGTCTGGGCT GGGGGGGGACA CTGTCCAAGG GAGTGGCCCC TATGAGTTTA TATTTTAACC   1740
        ACTGCTTCAA ATCTCGATTT CACTTTTTTT ATTTATCCAG TTATATCTAC ATATCTGTCA   1800
        TCTAAATAAA TGGCTTTCAA ACAAAGCAAC TGGGTCATTA AAACCAGCTC AAAAGGGGGTT   1860
        TAAAAAAAAA AAAACCAGCC CATCCTTTGA GGCTGATTTT TCTTTTTTTT AAGTTCTATT   1920
        TTAAAAGCTA TCAAACAGCG ACATAGCCAT ACATCTGACT GCCTGACATG GACTCCTGCC   1980
        CACTTGGGGG AAACCTTATA CCCAGAGGAA AATACACACC TGGGGAGTAC ATTTGACAAA   2040
        TTTCCCTTAG GATTTCGTTA TCTCACCTTG ACCCTCAGCC AAGATTGGTA AAGCTGCGTC   2100
        CTGGCGATTC CAGGAGACCC AGCTGGAAAC CTGGCTTCTC CATGTGAGGG GATGGGAAAG   2160
        GAAAGAAGAG AATGAAGACT ACTTAGTAAT TCCCATCAGG AAATGCTGAC CTTTTACATA   2220
        AAATCAAGGA GACTGCTGAA AATCTCTAAG GGACAGGATT TTCCAGATCC TAATTGCAAA   2280
        TTTAGCAATA AGGAGAGGAG TCCAGGGGGA CAAATAAAGG CAGAGAGAGA GAGAGAGAGA   2340
        GGGAGAGGAA GAAAAGAGAG AGAGAAAAGA GCCTCGTGCC
```

Seq ID NO: 100 Protein sequence
Protein Accession #: NP_003705

```
        1          11         21         31         41    .    51
        |          |          |          |          |          |
        MCAERLGQFM TLALVLATFD PARGTDATNP PEGPQDRSSQ QKGRLSLQNT AEIQHCLVNA   60
        GDVGCGVFEC FENNSCEIRG LHGICMTFLH NAGKFDAQGK SFIKDALKCK AHALRHRFGC   120
        ISRKCPAIRE MVSQLQRECY LKHDLCAAAQ ENTEVIVEMI HFKDLLLHEP YVDLVNLLLT   180
        CGEEVKEAIT HSVQVQCEQN NGSLCSILSF CTSAIQKPPT APPERQPQVD RTKLSRARHG   240
        EAGHHLPEPS SRETGRGAKG ERGSKSRPNA HARGRVGGLG AQGPSGSSEW EDEQSEYSDI   300
        RR
```

Seq ID NO: 101 DNA sequence
Nucleic Acid Accession #: NM_005940
Coding sequence: 23..1489

```
        1          11         21         31         41         51
        |          |          |          |          |          |
```

```
AAGCCCAGCA GCCCCGGGGC GGATGGCTCC GGCGCCTGG CTCCGCAGCG CGGCCGCGCG    60
CGCCCTCCTG CCCCCGATGC TGCTGCTGCT GCTCCAGCCG CCGCCGCTGC TGGCCCGGGC   120
TCTGCCGGCCG GACGTCCACC ACCTCCATGC CGAGAGGAGG GGGCCACAGC CCTGGCATGC   180
AGCCCTGCCC AGTAGCCCGG CACCTGCCCC TGCCACGCAG GAAGCCCCCC GGCCTGCCAG   240
CAGCCTCAGG CCTCCCCGCT GTGGCGTGCC CGACCCATCT GATGGGCTGA GTGCCCGCAA   300
CCGACAGAAG AGGTTCGTGC TTTCTGGCGG GCGCTGGGAG AAGACGGACC TCACCTACAG   360
GATCCTTCGG TTCCCCATGC AGTTGGTGCA GGAGCAGGTG CGGCAGACGA TGGCAGAGGC   420
CCTAAAGGTA TGGGAGCGATG TGACGCCACT CACCTTTACT GAGGTGCACG AGGGCCGTGC   480
TGACATCATG ATCGACTTCG CCAGGTACTG GCATGGGGAC GACCTGCCGT TTGATGGGCC   540
TGGGGGCATC CTGGCCCATG CCTTCTTCCC CAAGACTCAC CGAGAAGGGG ATGTCCACTT   600
CGACTATGAT GAGACCTGGA CTATCGGGGA TGACCAGGGC ACAGACCTGC TGCAGGTGGC   660
AGCCCATGAA TTTGGCCACG TGCTGGGGCT GCAGCACACA ACAGCAGCCA AGGCCCTGAT   720
GTCCGCCTTC TACACCTTTC GCTACCCACT GAGTCTCAGC CCAGATGACT GCAGGGGCGT   780
TCAACACCTA TATGGCCAGC CCTGGCCCAC TGTCACCTCC AGGACCCCAG CCCTGGGCCC   840
CCAGGCTGGG ATAGACACCA ATGAGATTGC ACCGCTGGAG CCAGACGCCC CGCCAGATGC   900
CTGTGAGGCC TCCTTTGACG CGGTCTCCAC CATCCGAGGC GAGCTCTTTT TCTTCAAAGC   960
GGGCTTTGTG TGGCGCCTCC GTGGGGGCCA GCTGCAGCCC GGCTACCCAG CATTGGCCTC  1020
TCGCCACTGG CAGGGACTGC CCAGCCCTGT GGACGCTGCC TTCGAGGATG CCCAGGGCCA  1080
CATTTGGTTC TTCCAAGGTG CTCAGTACTG GGTGTACGAC GGTGAAAAGC CAGTCCTGGG  1140
CCCCGCACCC CTCACCGAGC TGGGCCTGGT GAGGTTCCCG GTCCATGCTG CCTTGGTCTG  1200
GGGTCCCGAG AAGAACAAGA TCTACTTCTT CCGAGGCAGG GACTACTGGC GTTTCCACCC  1260
CAGCACCCGG CGTGTAGACA GTCCCGTGCC CCGCAGGGCC ACTGACTGGA GAGGGGTGCC  1320
CTCTCGAGATC GACGCTGCCT TCCAGGATGC TGATGGCTAT GCCTACTTCC TGCGCGGGCTT  1380
CCTCTACTGG AAGTTTGACC CTGTGAAGGT GAAGGCTCTG GAAGGCTTCC CCCGTCTCGT  1440
GGGTCCTGAC TTCTTTGGCT GTGCCGAGCC TGCCAACACT TTCCTCTGAC CATGGCTTGG  1500
ATGCCCTCAG GGGTGCTGAC CCCTGCCAGG CCACGAATAT CAGGCTAGAG ACCCATGGCC  1560
ATCTTTGTGG CTGTGGGCAC CAGGCATGGG ACTGAGCCCA TGTCTCCTGC AGGGGGATGG  1620
GGTGGGGTAC AACCACCATG ACAACTGCCG GGAGGGCCAC GCAGGTCGTG GTCACCTGCC  1680
AGCGACTGTC TCAGACTGGG CAGGGAGGCT TTGGCATGAC TTAAGAGGAA GGGCAGTCTT  1740
GGGACCCGCT ATGCAGGTCC TGGCAAACCT GGCTGCCCTG TCTCATCCCT GTCCCTCAGG  1800
GTAGCACCAT GGCAGGACTG GGGGAACTGG AGTGTCCTTG CTGTATCCCT GTTGTGAGGT  1860
TCCTTCCAGG GGCTGGCACT GAAGCAAGGG TGCTGGGGGC CCATGGCCTT CAGCCCTGGC  1920
TGAGCAACTG GGCTGTAGGG CAGGGCCACT TCCTGAGGTC AGGTCTTGGT AGGTGCCTGC  1980
ATCTGTCTGC CTTCTGGCTG ACAATCCTGG AAATCTGTTC TCCAGAATCC AGGCCAAAAA  2040
GTTCACAGTC AAATGGGGAG GGGTATTCTT CATGCAGGAG ACCCCAGGCC CTGGAGGCTG  2100
CAACATACCT CAATCCTGTC CCAGGCCGGA TCCTCCTGAA GCCCTTTTCG CAGCACTGCT  2160
ATCCTCCAAA GCCATTGTAA ATGTGTGTAC AGTGTGTATA AACCTTCTTC TTCTTTTTTT  2220
TTTTTAAACT GAGGGATTGTC ATTAAACACA GTTGTTTTCT
```

Seq ID NO: 102 **Protein sequence**
Protein Accession #: NP_005931

```
1          11         21         31         41         51
|          |          |          |          |          |
MAPAAWLRSA AARALLPPML LLLLQPPPLL ARALPPDVHH LHAERRGPQP WHAALPSSPA    60
PAPATQEAPR PASSLRPPRC GVPDPSDGLS ARNRQKRPVL SGGRWEKTDL TYRILRFPWQ   120
LVQEQVRQTM AEALKVWSDV TPLTFTEVHE GRADIMIDFA RYWHGDDLPF DGPGGILAHA   180
FFPKTHREGD VHFDYDETWT IGDDQGTDLL QVAAHEFGHV LGLQHTTAAK ALMSAFYTFR   240
YPLSLSFDDC RGVQHLYGQP WPTVTSRTPA LGPQAGIDTN EIAPLEPDAP PDACEAEFDA   300
VSTIRGELFF FKAGFVWRLR GGQLQPGYPA LASRHWQGLP SPVDAAFEDA QGHIWFFQGA   360
QYWVYDGEKP VLGPAPLTEL GLVRFPVHAA LVWGPEKNKI YPFRGRDYWR FHPSTRRVDS   420
PVPRRAIDWR GVPSEIDAAF QDADGYAYFL RGRLYWKFDP VKVKALEGFP RLVGPDFFGC   480
AEPANTFL
```

Seq ID NO: 103 **DNA sequence**
Nucleic Acid Accession #: NM_033151.2
Coding sequence: 351..4499

```
1          11         21         31         41         51
|          |          |          |          |          |
ACTGGGATAA AGCAAGAAGA CTGATTTTAT GAGCAGGGGT TTGATACATC AAAGGAGATT    60
GCCCAGGATC AAGGGTGCGG TGTTGGGGGT GGGTTGGGGA GGGTGGTTAG AGAAGGTTTC   120
ACTAAGTGAT TTGGGCCTGA GGCCTGAGAA GATGTTTAAA AAGAGGGATC AAGCACAGGC   180
TAAGGAGAGG AAAGAGCAGG CACCCAAACC TCTGCATGGC CCCAATATGC TCCCTGCAGG   240
GTAGTGCCCC CTCTTCTGGC TGCTCAAGGC GAGATCTAAG CTTCTTCTAA CTCCTGCTGT   300
CTTTTCATAT TCTCTGATTC TGGGAAACGA AGAATTGGCA GGAACTGAAA ATGACTAGGA   360
AGAGGACATA CTGGGTGCCC AACTCTTCTG GTGGCCTCGT GAATCGTGGC ATCGACATAG   420
GCGATGACAT GGTTTCAGGA CTTATTTATA AAACCTATAC TCTCCAAGAT GGCCCCTGGA   480
GTCAGCAAGA GAGAAATCCT GAGGCTCCAG GGAGGGCAGC TGTCCCACCG TGGGGGAAGT   540
ATGATGCTGC CTTGAGAACC ATGATTCCCT TCCGTCCCAA GCCGAGGTTT CCTGCCCCCC   600
AGCCCTGGA  CAATGCTGGC CTGTTCTCCT ACCTCACCGT GTCATGGCTC ACCCCGCTCA   660
TGATCCAAAG CTTACGGAGT CGCTTAGATG AGAACACCAT CCCTCCACTG TCAGTCCATG   720
ATGCCTCAGA CAAAAATGTC CAAAGGCTTC ACCGCCTTTG GGAAGAAGAA GTCTCAAGGC   780
GAGGGATTGA AAAAGCTTCA GTGCTTCTGG TGATGCTGAAG GTTCCAGAGA ACAAGGTTGA   840
TTTTCGATGC ACTTCTGGGC ATCTGCTTCT GCATTGCCAG TGTACTCGGG CCAATATTGA   900
TTATACCAAA GATCCTGGAA TATTCAGAAG AGCAGTTGGG GAATGTTGTC CATGGAGTGG   960
GACTCTGCTT TGCCCTTTTT CTCTCCGAAT GTGTGAAGTC TCTGAGTTTC TCCTCCAGTT  1020
GGATCATCAA CCAACGCACA GCCATCAGGT TCCGAGCAGC TGTTTCCTCC TTTGCCTTTG  1080
AGAAGCTCAT CCAATTTAAG TCTGTAATAC ACATCACCTC AGGAGAGGCC ATCAGCTTCT  1140
TCACCGGTGA TGTAAACTAC CTGTTTGAAG GGGTGTGCTA TGGACCCCTA GTACTGATCA  1200
CCTGCGCATC GCTGGTCATC TGCAGCATTT CTTCCTACTT CATTATTGGA TACACTGCAT  1260
TTATTGCCAT CTTATGCTAT CTCCTGGTTT TCCCACTGGC GGTATTCATG ACAAGAATGG  1320
CTGTGAAGGC TCAGCATCAC ACATCTGAGG TCAGCGACCA GCGCATCCGT GTGACCAGTG  1380
AAGTTCTCAC TTGCATTAAG CTGATTAAAA TGTACACATG GGAGAAACCA TTTGCAAAAA  1440
```

```
     TCATTGAAGA CCTAAGAAGG AAGGAAAGGA AACTATTGGA GAAGTGCGGG CTTGTCCAGA 1500
     GCCTGACAAG TATAACCTTG TTCATCATCC CCACAGTGGC CACAGCGGTC TGGGTTCTCA 1560
     TCCACACATC CTTAAAGCTG AAACTCACAG CGTCAATGGC CTTCAGCATG CTGGCCTCCT 1620
     TGAATCTCCT TCGGCTGTCA GTGTTCTTTG TGCCTATTGC AGTCAAAGGT CTCACGAATT 1680
     CCAAGTCTGC AGTGATGAGG TTCAAGAAGT TTTTCCTCCA GGAGAGCCCT GTTTTCTATG 1740
     TCCAGACATT ACAAGACCCC AGCAAAGCTC TGGTCTTTGA GGAGGCCACC TTGTCATGGC 1800
     AACAGACCTG TCCCGGGATC GTCAATGGGG CACTGGAGCT GGAGAGGAAC GGGGCATGCTT 1860
     CTGAGGGGAT GACCAGGCCT AGAGATGGCC TCGGGCCAGA GGAAGAAGGG AACAGCCTGG 1920
     GCCCAGAGTT GCACAAGATC AACCTGGTGG TGTCCAAGGG GATGATGTTA GGGGTCTGCG 1980
     GCAACACGGG GAGTGGTAAG AGCAGCCTGT TGTCAGCCAT CCTGGAGGAG ATGCACTTGC 2040
     TCGAGGGCTC GGTGGGGGTG CAGGGAAGCC TGGCCTATGT CCCCCAGCAG GCCTGGATCG 2100
     TCAGCGGGAA CATCAGGGAG AACATCCTCA TGGGGAGGCGC ATATGACAAG GCCCGATACC 2160
     TCCAGGTGCT CCACTGCTGC TCCCTGAATC GGGACCTGGA ACTTCTGCCC TTTGGGAGACA 2220
     TGACAGAGAT TGGAGAGCGG GGCCTCAACC TCTCTGGGGG GCAGAAACAG AGGATCAGCC 2280
     TGGCCCGCGC CGTCTATTCC GACCGTCAGA TCTACCTGCT GGACGACCCC CTGTCTGCTG 2340
     TGGACGCCCA CGTGGGGAAG CACATTTTTG AGGAGTGCAT TAAGAAGACA CTCAGGGGGA 2400
     AGACGGTCGT CCTGGTGACC CACCAGCTGC AGTACTTAGA ATTTTGTGGC CAGATCATTT 2460
     TGTTGGAAAA TGGGAAAATC TGTGAAAATG GAACTCACAG TGAGTTAATG CAGAAAAAGG 2520
     GGAAATATGC CCAACTTATC CAGAAGATGC ACAAGGAAGC CACTTCGGAC ATGTTGCAGG 2580
     ACACAGCAAA GATAGCAGAG AAGCCAAAGG TAGAAAGTCA GGCTCTGGCC ACCTCCCTGG 2640
     AAGAGTCTCT CAACGGAAAT GCTGTGCCGG AGCATGAGGT CACACAGGAG GAGGAGATGG 2700
     AAGAAGGCTC CTTGAGTTGG AGGGTCTACC ACCACTACAT CCAGGCAGCT GGAGGTTACA 2760
     TGGTCTCTTG CATAATTTTC TTCTTCGTGG TGCTGATCGT CTTCTTAACG ATCTTCAGCT 2820
     TCTGGTGGCT GAGCTACTGG TTGGAGCAGG GCTCGGGGAC CAATAGCAGC CGAGAGAGCA 2880
     ATGGAACCAT GGCAGACCTG GGCAACATTG CAGACAATCC TCAACTGTCC TTCTACCAGC 2940
     TGGTGTACGG GCTCAACGCC CTGCTCCTCA TCTGTGTGGG GGTCTGCTCC TCAGGGATTT 3000
     TCACCAAAGT CACGAGGAAG GCATCCACGG CCCTGCCACA CAAGCTCTTC AACAAGGTTT 3060
     TCCGCTGCCC CATGAGTTTC TTTGACACCA TCCCAATAGG CCGGCTTTTG AACTGCTTCG 3120
     CAGGGGACTT GGAACAGCTG GACCAGCTCT TGCCCATCTT TTCAGAGCAG TTCCTGGTCC 3180
     TGTCCTTAAT GGTGATCGCC GTCCTGTTGA TTGTCAGTGT GCTGTCTCCA TATATCCGGT 3240
     TAATGGGAGC CATAATCATG GTTATTTGCT TCATTTATTA TATGATGTTC AAGAAGGCCA 3300
     TCGGTGTGTT CAAGAGACTG GAGAACTATA GCCGGTCTCC TTTATTCTCC CACATCCTCA 3360
     ATTCTCTGCA AGGCCTGAGC TCCATCCATG TCTATGGAAA AACTGAAGAC TTCATCAGCC 3420
     AGTTTAAGAG GCTGACTGAT GCGCAGAATA ACTACCTGCT GTTGTTCTA TCTTCCACAC 3480
     GATGGATGGC ATTGAGGCTG GAGATCATGA CCAACCTTGT GACCTTGGCT GTTGCCCTGT 3540
     TCGTGGCTTT TGGCATTTCC TCCACCCCCT ACTCCTTTAA AGTCATGGCT GTCAACATCG 3600
     TGCTGCAGCT GGCGTCCAGC TTCCAGGCCA CTGCCCGGAT TGGCTTGGAG ACAGAGGCAC 3660
     AGTTCACGGC TGTAGAGAGG ATACTGCAGT ACATGAAGAT GTGTGTCTCG GAAGCTCCTT 3720
     TACACATGGA AGGCACAAGT TGTCCCCAGG GGTGGCCCCA GCATGGGGAA ATCATATTTC 3780
     AGGATTATCA CATGAAATAC AGAGACAACA CACCCACCGT GCTTCACGGC ATCAACCTGA 3840
     CCATCCGCGG CCACGAAGTG GTGGGGCATCG TGGGAAGGAC GGGGCTCTGGG AAGTCCTCCT 3900
     TGGGCATGGC TCTCTTCCGC CTGGTGGAGC CCATGGCAGG CCGGATTCTC ATTGACGGCG 3960
     TGGACATTTG CAGCATCGGC CTGGAGGACT TGCGGTCAGA GCTCTCAGTG ATCCCTCAAG 4020
     ATCCAGTGCT GCTCTCAGGA ACCATCAGAT TCAACCTAGA TCCCTTTGAC CGTCACACTG 4080
     ACCAGCAGAT CTGGGATGCC TTGGAGAGGA CATTCCTGAC CAAGGCCATC TCAAAGTTCC 4140
     CCAAAAAGCT GCATACAGAT GTGGTGGAAA ACGGTGGAAA CTTCTCTGTG GGGGAGAGGC 4200
     AGCTGCTCTG CATTGCCAGG GCTGTGCTTC GCAACTCCAA GATCATCCTT ATCGATGAAG 4260
     CCACAGCCTC CATTGACATG GAGACAGACA CCCTGATCCA GCGCACAATC CGTGAAGCCT 4320
     TCCAGGGCTG CACCGTGCTC GTCATTGCCC ACCGTGTCAC CACTGTGCTG AACTGTGAAT 4380
     ACATCCTGGT TATGGGCAAT GGGAAGGTGG TAGAATTTGA TCGGCCGGAG GTACTGCGGA 4440
     AGAAGCCTGG GTCATTGTTC GCAGCCCTCA TGGCCACAGC CACTTCTTCA CTGAGATAAG 4500
     GAGATGTGGA GACTTCATGG AGGCTGGCAG CTGAGCTCAG AGGTTCACAC AGGTGCAGGT 4560
     TCGAGGCCCA CAGTCTGCGA CCTTCTTGTT TGGAGATGAG AACTTCTCCT GGAAGCAGGG 4620
     GTAAATGTAG GGGGGGTGGG GATTGCTGGA TGGAAACCCT GGAATAGGCT ACTTGATGGC 4680
     TCTCAAGACC TTAGAACCCC AGAACCATCT AAGACATGGA ATTCAGTGAT CATGTGGTTC 4740
     TCCTTTTAAC TTACATGCTG AATAATTTTA TAATAAGGTA AAAGCTTATA GTTTTCTGAT 4800
     CTGTGTTAGA AGTGTTGCAA ATGCTGTACT GACTTTGTAA AATATAAAAC TAAGGAAAAC 4860
     TC
```

Seq ID NO: 104 Protein sequence
Protein Accession #: NP_149163.2

```
     1          11          21          31          41          51
     |          |           |           |           |           |
     MTRKRTYWVP NSSGGLVNRG IDIGDDMVSG LIYKTYTLQD GPWSQQERNP EAPGRAAVPP    60
     WGKYDAALRT MIPFRPKPRF PAPQPLDNAG LPSYIAVSWL TPIMIQSLRS RLDENTIPPL   120
     SVHDASDKNV QRLHRLWEEE VSRRGIEKAS VLLVMLRFQR TRLIFDALLG ICFCIASVLG   180
     PILIIPKILE YSEEQLGNVV HGVGLCFALF LSECVKSLSF SSSWIINQRT AIRFRAAVSS   240
     FAFEKLIQPK SVIHITSGEA ISFFTGDVNY LFEGVCYGPL VLITCASLVI CSISSYFIIG   300
     YTAFIAILCY LLVFPLAVFM TRMAVKAQHH TSEVSDQRIR VTSEVLTCIK LIKMYTWEKP   360
     FAKIIEDLRK KERKLLEKCG LVQSLTSITL FIIPTVATAV WVLIRTSLKL KLTASMAPSM   420
     LASLNLLRLS VFFVPIAVKG LTNSKSAVMR FKKFFLQESP VFYVQTLQDP SKALVFEEAT   480
     LSWQQTCPGI VNGALELERN GHASEGMTRP RDALGPEEEG NSLGPELHKI NLVVSKGMML   540
     GVCGNTGSGK SSLLSAILEE MHLLEGSVGV QGSLAYVPQQ AWIVSGNIRE NILMGGAYDK   600
     ARYLQVLHCC SLNRDLELLP FGDMTEIGER GLNLSGGQKQ RISLARAVYS DRQIYLLDDP   660
     LSAVDAHVGK HIFEECIKKT LRGKTVVLVT EQLQYLEFCG QIILLENGKI CENGTHSELM   720
     QRKGKYAQLI QKMHKEATSD MLQDTAKIAE KPKVESQALA TSLEESLNGN AVPEHQLTQE   780
     EEMEEGSLSW RVYHHYIQAA GGYMVSCIIP FFVVLIVFLT IPSPWWLSYW LEQGSGTNSS   840
     RESNGTMADL GNIADNPQLS FYQLVYGLNA LLLICVGVCS SGIPTKVTRK ASTALHNKLF   900
     NKVFRCPMSF FDTIPIGRLL NCFAGDLEQL DQLLPIFSEQ FLVLSLMVIA VLLIVSVLSP   960
     YILLMGAIIM VICFIYYMMF KKAIGVFKRL ENYSRSPLFS HILNSLQGLS SIHVYGKTED  1020
     FISQFKRLTD AQNNYLLLFL SSTRWMALRL EIMTNLVTLA VALFVAFGIS STPYSFKVMA  1080
     VNIVLQLASS FQATARIGLE TEAQPTAVER ILQYMKMCVS EAPLHMEGTS CPQGWPQHGE  1140
     IIFQDYHMKY RDNTPTVLHG INLTIRGHEV VGIVGRTGSG KSSLGMALFR LVEPMAGRIL  1200
```

IDGVDICSIG LEDLRSKLSV IPQDPVLLSG TIRFNLDPFD RHTDQQIWDA LERTFLTKAI 1260
SKFPKKLHID VVENGGNPSV GERQLLCIAR AVLRNSKIIL IDEATASIDM ETDTLIQRTI 1320
REAFQGCTVL VIAHRVTTVL NCDHILVMGN GKVVEFDRPE VLRKKPGSLF AALMATATSS 1380
LR

Seq ID NO: 105 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..4043

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGACTAGGA AGAGGACATA CTGGGTGCCC AACTCTTCTG GTGGCCTCGT GAATCGTGGC   60
ATCGACATAG GCGATGACAT GGTTTCAGGA CTTATTTATA AAACCTATAC TCTCCAAGAT  120
GGGCCCTGGA GTCAGCAAGA GAGAAATCCT GAGGCTCCAG GGAGGGCAGC TGTCCCACCG  180
TGGGGGAAGT ATGATGCTGC CTTGAGAACC ATGATTCCCT TCCGTCCCAA GCCGAGGTTT  240
CCTGCCCCCC AGCCCCTGGA CAATGCTGGC CTGTTCTCCT ACCTCACCGT GTCATGGCTC  300
ACCCCGCTCA TGATCCAAAG CTTACGGAGT CGCTTAGATG AGAACACCAT CCCTCCACTG  360
TCAGTCCATG ATGCCTCAGA CAAAAATGTC CAAAGGCTTC ACCGCCTTTG GGAAGAAGAA  420
GTCTCAAGGC GAGGGATTGA AAAAGCTTCA GTGCTTCTGG TGATGCTGAG GTTCCAGAGA  480
ACAAGGTTGA TTTTCGATGC ACTTCTGGGC ATCTGCTTCT GCATTGCCAG TGTACTCGGG  540
CCAATATTGA TTATACCAAA GATCCTGGAA TATTCAGAAG AGCAGTTGGG GAATGTTGTC  600
CATGGAGTGG GACTCTGCTT TGCCCTTTTT CTCTCCGAAT GTGTGAAGTC TCTGAGTTTC  660
TCCTCCAGTT GGATCATCAA CCAACGCACA GCCATCAGGT TCCGAGCAGC TGTTTCCTCC  720
TTTGCCTTTG AGAAGCTCAT CCAATTTAAG TCTGTAATAC ACATCACCTC AGGAGAGGCC  780
ATCAGCTTCT TCACCGGTGA TGTAAACTAC CTGTTTGAAG GGGTGTGCTA TGGACCCCTA  840
GTACTGATCA CCTGCGCATC GCTGGTCATC TGCAGCATTT CTTCCTACTT CATTATTGGA  900
TACACTGCAT TTATTGCCAT CTTATGCTAT CTCCTGGTTT TCCCACTGGC GGTATTCATG  960
ACAAGAATGG CTGTGAAGGC TCAGCATCAC ACATCTGAGG TCAGCGACCA GCGCATCCGT 1020
GTGACCAGTG AAGTTCTCAC TTGCATTAAG CTGATTAAAA TGTACACATG GGAGAAACCA 1080
TTTGCAAAAA TCATTGAAGG TATGGAAAGT CTGACTTTCT GCTCCAAACC TGGTGATGGC 1140
ATGGCCTTCA GCATGCTGGC CTCCTTGAAT CTCCTTCGGC TGTCAGTGTT CTTTGTGCCT 1200
ATTGCAGTCA AAGGTCTCAC GAATTCCAAG TCTGCAGTGA TGAGGTTCAA GAAGTTTTTC 1260
CTCCAGGAGA GCCCTGTTTT CTATGTCCAG ACATTACAAG ACCCCAGCAA AGCTCTGGTC 1320
TTTGAGGAGG CCACCTTGTC ATGGCAACAG ACCTGTCCCG GGATCGTCAA TGGGGCACTG 1380
GAGCTGGAGA GGAACGGGCA TGCTTCTGAG GGGATGACCA GGCCTAGAGA TGCCCTCGGG 1440
CCAGAGGGAG AAGGGAACAG CCTGGGCCCA GAGTTGCACA AGATCAACCT GGTGGTGTCC 1500
AAGGGGATGA TGTTAGGGGT CTGCGGCAAC ACGGGGAGTG GTAAGAGCAG CCTGTTGTCA 1560
GCCATCCTGG AGGAGATGCA CTTGCTCGAG GGCTCGGTGG GGGTGCAGGG AAGCCTGGCC 1620
TATGTCCCCC AGCAGGCCTG GATCGTCAGC GGGAACATCA GGGAGAACAT CCTCATGGGA 1680
GGGCGCATATG ACAAGGCCCG ATACCTCCAG GTGCTCCACT GCTGCTCCCT GAATCGGGAC 1740
CTGGAACTTC TGCCCTTTGG AGACATGACA GAGATTGGAG AGCGGGGCCT CAACCTCTCT 1800
GGGGGGCAGA AACAGAGGAT CAGCCTGGCC CGCGCCGTCT ATTCCGACCG TCAGATCTAC 1860
CTGCTGGACG ACCCCCTGTC TGCTGTGGAC GCCCACGTGG GGAAGCACAT TTTTGAGGAG 1920
TGCATTAAGA AGACACTCAG GGGGAAGACG GTCGTCCTGG TGACCCACCA GCTGCAGTAC 1980
TTAGAATTTT GTGGCCAGAT CATTTTGTTG GAAAATGGGA AAATCTGTGA AAATGGAACT 2040
CACAGTGAGT TAATGCAGAA AAAGGGGAAA TATGCCCAAC TTATCCAGAA GATGCACCAG 2100
GAAGCCACTT CGGACATGTT GCAGGACACA GCAAAGATAG CAGAGAAGCC AAAGGTAGAA 2160
AGTCAGGCTC TGGCCACCTC CCTGGAAGAG TCTCTCAACG GAAATGCTGT GCCGGAGCAT 2220
CAGCTCACAC AGGAGGAGGA GATGGAAGAA GGCTCCTTGA GTTGGAGGGT CTACCACCAC 2280
TACATCCAGG CAGCTGGAGG TTACATGGTC TCTTGCATAA TTTTCTTCTT CGTGGTGCTG 2340
ATCGTCTTCT TAACGATCTT CAGCTTCTGG TGGCTGAGCT ACTGGTTGGA GCAGGGCTCG 2400
GGGACCAATA GCAGCCGAGA GAGCAATGGA ACCATGGCAG ACCTGGGCAA CATTGCAGAC 2460
AATCCTCAAC TGTCCTTCTA CCAGCTGGTG TACGGGCTCA ACGCCCTGCT CCTCATCTGT 2520
GTGGGGGTCT GCTCCTCAGG GATTTTCACC AAAGTCACGA GGAAGGCATC CACGGCCCTG 2580
CACAACAAGC TCTTCAACAA GGTTTTCCGC TGCCCCATGA GTTTCTTTGA CACCATCCCA 2640
ATAGGCCGGC TTTTGAACTG CTTCGCAGGG GACTTGGAAC AGCTGGACCA GCTCTTGCCC 2700
ATCTTTTCAG AGCAGTTCCT GGTCCTGTCC TTAATGGTGA TCGCCGTCCT GTTGATTGTC 2760
AGTGTGCTGT CTCCATATAT CCTGTTAATG GGAGCCATAA TCATGGTTAT TTGCTTCATT 2820
TATTATATGA TGTTCAAGAA GGGCCATCGGT GTGTTCAAGA CTATAGCCGG 2880
TCTCCTTTAT TCTCCCACAT CCTCAATTCC CTGCAAGGCC TGAGCTCCAT CCATGTCTAT 2940
GGAAAAACTG AAGACTTCAT CAGCCCAGTTT AAGAGGCTGA CTGATGCGCA GAATAACTAC 3000
CTGCTGTTCT TTCTATCTTC CACACGATGG ATGGCATTGA GGCTGGAGAT CATGACCAAC 3060
CTTGTGACCT TGGCTGTTGC CCTGTTCGTG GCTTTTGGCA TTTCCTCCAC CCCCTACTCC 3120
TTTAAAGTCA TGGCTGTCAA CATCGTGCTG CAGCTGGCGT CCAGCTTCCA GGCCACTGCC 3180
CGGATTGGCT TGGAGACAGA GGCACAGTTC ACGGCTGTAG AGAGGGATACT GCAGTACATG 3240
AAGATGTGTG TCTCGGAAGC TCCTTTACAC ATGGAAGGCA CAAGTTGTCC CCAGGGGTGG 3300
CCACAGCATG GGGAAATCAT ATTTCAGGAT TATCACATGA AATACAGAGA CAACACACCC 3360
ACGGTGCTTC ACGGCATCAA CCTGACCATC CGCGGCCACG AAGTGGTGGG CATCGTGGGA 3420
AGGACGGGCT CTGGGAAGTC CTCCTTGGGC ATGGCTCTCT TCCGCCTGGT GGAGCCCATG 3480
GCAGGCCGGA TTCTCATTGA CGGCGTGGAC ATTTGCAGCA TCGGCCTGGA GGACTTGCGG 3540
TCCAAGCTCT CAGTGATCCC TCAAGATCCA GTGCTGCTCT CAGGAACCAT CAGATTCAAC 3600
CTAGATCCCT TTGACCGTCA CACTGACCAG CAGATCTGGG ATGCCTTGGA GAGGACATTC 3660
CTGACCAAGG CCATCTCAAA GTTCCCCAAA AAGCTGCATA CAGATGTGGT GGAAAACGGT 3720
GGAAACTTCT CTGTGGGGGA GAGGCAGCTG CTCTGCATTG CCAGGGCTGT GCTTCGCAAC 3780
TCCAAGATCA TCCTTATCGA TGAAGCCACA GCCTCCATTG ACATGGAGAC AGACACCCTG 3840
ATCCAGCGCA CAATCCGTGA AGCCTTCCAG GGCTGCACCG TGCTCGTCAT TGCCCACCGT 3900
GTCACCACTG TGCTGAACTG TGACCACATC CTGGTTATGG CAATGGGAA GGTGGTAGAA 3960
TTTGATCGGC CGGAGGTACT GCGGAAGAAG CCTGGGTCAT TGTTCGCAGC CCTCATGGCC 4020
ACAGCCACTT CTTCACTGAG ATAA
```

Seq ID NO: 106 Protein sequence
Protein Accession #: Eos sequence

```
          1         11        21        31        41        51
          |         |         |         |         |         |
     MTRKRTYWVP NSSGGLVNRG IDIGDDMVSG LIYKTYTLQD GPWSQQERNP EAPGRAAVPP    60
     WGKYDAALRT MIPFRPKPRP PAPQPLDNAG LPSYLTVSWL TPLMIQSLRS RLDENTIPPL'  120
     SVHDASDKNV QRLHRLWEES VSRRGIEKAS VLLVMLRFQR TRLIFDALLG ICFCIASVLG   180
     PILIIPKILE YSEEQLGNVV HGVGLCFALP LSECVKSLSP SSSWIINQRT AIRFRAAVSS   240
     FAFEKLIQFK SVIHITSGEA ISPPTGDVNY LPEGVCYGPL VLITCASLVI CSISSYFIIG   300
     YTAFIAILCY LLVFPLAVFM TRMAVKAQHH TSEVSDQRIR VTSEVLTCIK LIKMYTWEKP   360
     PAKIIEGMES LTFCSKPGDG MAFSMLASLN LLRLSVPFVP IAVKGLTNSK SAVMRFKKFF   420
     LQESPVFYVQ TLQDPSKALV FEEATLSWQQ TCPGIVNGAL ELERNGHASE GMTRPRDALG   480
     PEEEGNSLGP ELEKINLVVS KGMMLGVCGN TGSGKSSLLS AILEEMHLLE GSVGVQGSLA   540
     YVPQQAWIVS GNIRENILMG GAYDKARYLQ VLHCCSLNRD LELLPFGDMT EIGERGLNLS   600
     GGQKQRISLA RAVYSDRQIY LLDDPLSAVD AHVGKHIFEE CIKKTLRGKT VVLVTHQLQY   660
     LRPCGQIILL ENGKICENGT HSELMQKKGK YAQLIQKMHK EATSDMLQDT AKIAEKPKVE   720
     SQALATSLEE SLNGNAVPEH QLTQEEEMEE GSLSWRVYHH YIQAAGGYMV SCIIFFPVVL   780
     IVFLTIPSFW WLSYWLEQGS GTNSSRESNG TMADLGNIAD NPQLSFYQLV YGLNALLLIC   840
     VGVCSSGIFT KVTRKASTAL HNKLFNKVFR CPMSFFDTIP IGRLLNCFAG DLEQLDQLLP   900
     IFSEQFLVLS LMVIAVLLIV SVLSPYILLM GAIIMVICFI YYMMFKKAIG VFKRLENYSR   960
     SPLFSHILNS LQGLSSIHVY GKTEDPISQF KRLTDAQNNY LLLFLSSTRW MALRLEIMTN  1020
     LVTLAVALFV AFGISSTPYS PKVMAVNIVL QLASSFQATA RIGLETEAQF TAVERILQYM  1080
     KMCVSEAPLH MEGTSCPQGW PQHGEIIFQD YHMKYRDNTP TVLHGINLTI RGHEVVGIVG  1140
     RTGSGKSSLG MALFRLVEPM AGRILIDGVD ICSIGLEDLR SKLSVIPQDP VLLSGTIRFN  1200
     LDPFDRHTDQ QIWDALERTF LTKAISKFPK KLHTDVVENG GNFSVGERQL LCIARAVLRN  1260
     SKIILIDEAT ASIDMETDTL IQRTIREAFQ GCTVLVIAHR VTTVLNCDHI LVMGNGKVVE  1320
     FDRPEVLRKK PGSLPAALMA TATSSLR
```

Seq ID NO: 107 DNA sequence
Nucleic Acid Accession #: NM_024022
Coding sequence: 202..1563

```
          1         11        21        31        41        51
          |         |         |         |         |         |
     ACCGGGCACC GGACGGCTCG GGTACTTTCG TTCTTAATTA GGTCATGCCC GTGTGAGCCA    60
     GGAAAGGGCT GTGTTTATGG GAAGCCAGTA ACACTGTGGC CTACTATCTC TTCCGTGGTG   120
     CCATCTACAT TTTTGGGACT CGGGAATTAT GAGGTAGAGG TGGAGGCGGA GCCGGGATGTC   180
     AGAGGTCCTG AAAATAGTCAC CATGGGGGAA AATGATCCGC CTGCTGTTGA AGCCCCCTTC   240
     TCATTCCGAT CGCTTTTTGG CCTTGATGAT TTGAAAATAA GTCCTGTTGC ACCAGATGCA   300
     GATGCTGTTG CTGCACAGAT CCTGTCACTG CTGCCATTGA AGTTTTTTCC AATCATCGTC   360
     ATTGGGATCA TTGCATTGAT ATTAGCACTG GCCATTGGTC TGGGCATCCA CTTCGACTGC   420
     TCAGGGAAGT ACAGATGTCG CTCATCCTTT AAGTGTATCG AGCTGATAGC TCGATGTGAC   480
     GGAGTCTCGG ATTGCAAAGA CGGGGAGGAC GAGTACCGCT GTGTCCGGGT GGGGTGGTCAG   540
     AATGCCGTGC TCCAGGTGTT CACAGCTGCT TCGTGGAAGA CCATGTGCTC CGATGACTGG ·  600
     AAGGGTCACT ACGCAAATGT TGCCTGTGCC CAACTGGGTT TCCCAAGCTA TGTGAGTTCA   660
     GATAACCTCA GAGTGAGCTC GCTGGAGGGG CAGTTCCGGG AGGAGTTTGT GTCCATCGAT   720
     CACCTCTTGC CAGATGACAA GGTGACTGCA TTACACCACT CAGTATATGT GAGGGAGGGA   780
     TGTGCCTCTG GCCACGTGGT TACCTTGCAG TGCACAGCCT GTGGTCATAG AAGGGGCTAC   840
     AGCTCACGCA TCGTGGGTGG AAACATGTCC TTGCTCTCGC AGTGGCCCTG GCAGGCCAGC   900
     CTTCAGTTCC AGGGCTACCA CCTGTGCGGG GGCTCTGTCA TCACGCCCCT GTGGATCATC   960
     ACTGCTGCAC ACTGTGTTTA TGACTTGTAC CTCCCCAAGT CATGGACCAT CCAGGTGGGT  1020
     CTAGTTTCCC TGTTGGACAA TCCAGCCCCA TCCCACTTGG TGGAGAAGAT TGTCTACCAC  1080
     AGCAAGTACA AGCCAAAGAG GCTGGGCAAT GACATCGCCC TTATGAAGCT GGCCGGGCCA  1140
     CTCACGGTCA ATGAAATGAT CCAGCCTGTG TGCCTGCCCA ACTCTGAAGA GAACTTCCCC  1200
     GATGGAAAAG TGTGCTGGAC GTCAGGATGG GGGGCCACAG AGGATGGAGG TGACGCCTCC  1260
     CCTGTCCTGA ACCACGCGGC CGTCCCTTTG ATTTCCAACA AGATCTGCAA CCACAGGGAC  1320
     GTGTACGGTG GCATCATCTC CCCCTCCATG CTCTGCGCGG GCTACCTGAC GGGTGGCGTG  1380
     GACAGCTGCC AGGGGGACAG CCGGGGGCCC CTGGTGTGTC AAGAGAGGAG GCTGTGGAAG  1440
     TTAGTGGGAG CGACCAGCTT TGGCATCGGC TGCCGCAGAGG TGAACAAGCC TGGGGTGTAC  1500
     ACCGTGTCA CCTCCTTCCT GGACTGGATC CACGAGCAGA TGGAGAGAGA CCTAAAAACC  1560
     TGAAGAGGAA GGGGACAAGT AGCCACCTGA GTTCCTGAGG TGATGAAGAC AGCCGGATCC  1620
     TCCCCTGGAC TCCGGTGTAG GAACCTGCAC ACGAGCAGAC ACCCTTGGAG CTCTGAGTTC  1680
     CGGCACCAGT AGCAGGCCCG AAAGAGGCAC CCTTCCATCT GATTCCAGCA CAACCTTCAA  1740
     GCTGCTTTTT GTTTTTTGTT TTTTTGAGGT GGAGTCTCGC TCTGTTGCCC AGGCTGGAGT  1800
     GCAGTGGCGA AATCCCTGCT CACTGCAGCC TCCGCTTCCC TGGTTCAAGC GATTCTCTTG  1860
     CCTCAGCTTC CCCAGTAGCT GGGACCACAG GTGCCCGCCA CCACACCCAA CTAATTTTTG  1920
     TATTTTTAGT AGAGACAGGG TTTCACCATG TTGGCCAGGC TGCTCTCAAA CCCCTGACCT  1980
     CAAATGATGT GCCTGCTTCA GCCTCCCACA GTGCTGGGAT TACAGGCATG GGCCACCACG  2040
     CCTAGCCTCA CGCTCCTTTC TGATCTTCAC TAAGAACAAA AGAAGCAGCA ACTTGCAAGG  2100
     GCGGCCTTTC CCACTGGTCC ATCTGGTTTT CTCTCCAGGG GTCTTGCAAA ATTCCTGACG  2160
     AGATAAGCAG TTATGTGACC TCACGTGCAA AGCCACCAAC AGCCACTCAG AAAAGACGCA  2220
     CCAGCCCAGA AGTGCAGAAC TGCAGTGACT GCACGTTTTC ATCTCTAGGG ACCAGAACCA  2280
     AACCCACCCT TTCTACTTCC AAGACTTATT TTCACATGTG GGGAGGTTAA TCTAGGAAATG  2340
     ACTCGTTTAA GGCCTATTTT CATGATTTCT TTGTAGCATT TGGTGCTTGA CGTATTATTG  2400
     TCCTTTGATT CCAAATAATA TGTTTCCTTC CCTCAAAAA AAAAAAAAAA AAAAAAAAAA  2460
     AAAAA
```

Seq ID NO: 108 Protein sequence
Protein Accession #: NP_076927

```
          1         11        21        31        41        51
          |         |         |         |         |         |
     MGEWDPPAVE APFSFRSLFG LDDLKISPVA PDADAVAAQI LSLLPLKFFP IIVIGIIALI    60
     LALAIGLGIH FDCSGKYRCR SSFKCIELIA RCDGVSDCKD GEDEYRCVRV GGQNAVLQVF   120
     TAASWKTMCS DDWRGHYANV ACAQLGFPSY VSSDWLRVSS LEGQPREEFV SIDHLLPDDK   180
     VTAIHHSVYV REGCASGHVV TLQCTACGHR RGYSSRIVGG NMSLLSQWPW QASLQFQGYH   240
```

```
LCGGSVITPL WIITAAHCVY DLYLPKSWTI QVGLVSLLDN PAPSHLVEKI VYHSKYKPKR   300
LGNDIALMKL AGPLTFNEMI QPVCLPNSEE NFPDGKVCWT SGWGATEDGG DASPVLNHAA   360
VPLISNKICN HRDVYGGIIS PSMLCAGYLT GGVDSCQGDS GGPLVCQERR LWKLVGATSF   420
GIGCAEVNKP GVYTRVTSFL DWIHEQMERD LKT
```

Seq ID NO: 109 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_000493.2
Coding sequence: 97..2139

```
1          11         21         31         41         51
|          |          |          |          |          |
CACCTTCTGC ACTGCTCATC TGGGCAGAGG AAGCTTCAGA AAGCTGCCAA GGCACCATCT    60
CCAGGAACTC CCAGCACGCA GAATCCATCT GAGAATATGC TGCCACAAAT ACCCTTTTTG   120
CTGCTAGTAT CCTTGAACTT GGTTCATGGA GTGTTTTACG CTGAACGATA CCAAATGCCC   180
ACAGGCATAA AAGGCCCACT ACCCAACACC AAGACACAGT TCTTCATTCC CTACACCATA   240
AAGAGTAAAG GTATAGCAGT AAGAGGAGAG CAAGGTACTC CTGGTCCACC AGGCCCTGCT   300
GGACCTCGAG GGCACCCAGG TCCTTCTGGA CCACCAGGAA AACCAGGCTA CGGAAGTCCT   360
GGACTCCAAG GAGAGCCAGG GTTGCCAGGA CCACCGGGAC CATCAGCTGT AGGGAAACCA   420
GGTGTGCCAG GACTCCCAGG AAAACCAGGA GAGAGAGGAC CATATGGACC AAAAGGAGAT   480
GTTGGACCAG CTGGCCTACC AGGACCCCGG GGCCCACCAG GACCACCTGG AATCCCTGGA   540
CCGGCTGGAA TTTCTGTGCC AGGAAAACCT GGACAACAGG GACCCACAGG AGCCCCAGGA   600
CCCAGGGGCT TTCCTGGAGA AAAGGGTGCA CCAGGAGTCC CTGGTATGAA TGGACAGAAA   660
GGGGAAATGG GATATGGTGC TCCTGGTCGT CCAGGTGAGA GGGGTCTTCC AGGCCCTCAG   720
GGTCCCACAG GACCATCTGG CCCTCCTGGA GTGGGAAAAA GAGGTGAAAA TGGGGTTCCA   780
GGACAGCCAG GCATCAAAGG TGATAGAGGT TTTCCGGGAG AAATGGGACC AATTGGCCCA   840
CCAGGTCCCC AAGGCCCTCC TGGGGAACGA GGGCCAGAAG GCATTGGAAA GCCAGGAGCT   900
GCTGGAGCCC CAGGCCAGCC AGGGATTCCA GGAACAAAAG GTCTCCCTGG GGCTCCAGGA   960
ATAGCTGGGC CCCCAGGGGC TCCTGGCTTT GGGAAACCAG GCTTGCCAGG CCTGAAGGGA  1020
GAAAGAGGAC CTGCTGGCCT TCCTGGGGGT CCAGGTGCCA AAGGGGAACA AGGGCCAGCA  1080
GGTCTTCCTG GGAAGCCAGG TCTGACTGGA CCCCCTGGGA ATATGGGACC CCAAGGACCA  1140
AAAGGCATCC CGGGTAGCCA TGGTCTCCCA GGCCCTAAAG GTGAGACAGG GCCAGCTGGG  1200
CCTGCAGGAT ACCCTGGGGC TAAGGGTGAA AGGGGTTCCC CTGGGTCAGA TGGAAAACCA  1260
GGGTACCCAG GAAAACCAGG TCTCGATGGT CCTAAGGGTA ACCCAGGGTT ACCAGGTCCA  1320
AAAGGTGATC CTGGAGTTGG AGGACCTCCT GGTCTCCCAG GCCCTGTGGG CCCAGCAGGA  1380
GCAAAGGGAA TGCCCGGACA CAATGGAGAG GCTGGCCCAA GAGGTGCCCC TGGAATACCA  1440
GGTACTAGAG GCCCTATTGG GCCACCAGGC ATTCCAGGAT TCCCTGGGTC TAAAGGGGAT  1500
CCAGGAAGTC CCGGTCCTCC TGGCCCAGCT GGCATAGCCA CTAAGGGCCT CAATGGACCC  1560
ACCGGGCCAC CAGGGCCTCC AGGTCCAAGA GGCCACTCTG GAGAGCCTGG TCTTCCAGGG  1620
CCCCCTGGGC CTCCAGGCCC ACCAGGTCAA GCAGTCATGC CTGGGGGTTT TATAAAGGCA  1680
GGGCAAAGGC CCAGTCTTTC TGGGACCCCT CTTGTTAGTG CCAACCAGGG GGTAACAGGA  1740
ATGCCTGTGT CTGCTTTTAC TGTTATTCTC TCCAAAGCTT ACCCAGCAAT AGGAACTCCC  1800
ATACCATTTG ATAAAATTTT GTATAACAGG CAACAGCATT ATGACCCAAG GACTGGAATC  1860
TTTACTTGTC AGATACCAGG AATATACTAT TTTTCATACC ACGTGCATGT GAAAGGGACT  1920
CATGTTTGGG TAGGCCTGTA TAAGAATGGC ACCCCTGTAA TGTACACCTA TGATGAATAC  1980
ACCAAAGGCT ACCTGGATCA GGCTTCAGGG AGTGCCATCA TCGATCTCAC AGAAAATGAC  2040
CAGGTGTGGC TCCAGCTTCC CAATGCCGAG TCAAATGACC TATACTCCTC TGAGTATGTC  2100
CACTCCTCTT TCTCAGGATT CCTAGTGGCT CCAATGTGAG TACACCCCAC AGAGCTAATC  2160
TAAATCTTGT GCTAGAAAAA GCATTCTCTA ACTCTACCCC ACCCTACAAA ATGCATATGG  2220
AGGTAGGCTG AAAAGAATGT AATTTTTATT TTCTGAAATA CAGATTTGAG CTATCAGACC  2280
AACAAACCTT CCCCCTGAAA AGTGAGCAGC AACGTAAAAA CGTATGTGAA GCCCTCTCTTG  2340
AATTTCTAGT TAGCAATCTT AAGGCTCTTT AAGGTTTTCT CCAATATTAA AAAATATCAC  2400
CAAAGAAGTC CTGCTATGTT AAAAACAAAC AACAAAAAAC AAAGCAACAA AAAAAAAAAC  2460
TAAAAAAAAA AACAGAAATA GAGCTCTAAG TTATGTGAAA TTTGATTTGA GAAACTCGGC  2520
ATTTCCTTTT TAAAAAAGCC TGTTTCTAAC TATGAATATG AGAACTTCTA GGAACATCC   2580
AGGAGGTATC ATATAACTTT GTAGAACTTA AATACTTGAA TATTCAAATT TAAAAGACAC  2640
TGTATCCCCT AAAATATTTC TGATGGTGCA CTACTCTGAG GCCTGTATGG CCCCTTTCAT  2700
CAATATCTAT TCAAATATAC AGGTGCATAT ATACTTGTTA AAGCTCTTAT ATAAAAAAGC  2760
CCCAAAAATAT TGAAGTTCAT CTGAAATGCA AGGTGCTTTC ATCAATGAAC CTTTTCAAAA  2820
CTTTTCATAG ATTGCAGAGA AGCTTTTTAT ATACCCAGCA TAACTTGGAA ACAGGTATCT  2880
GACCTATTCT TATTTAGTTA ACACAAGTGT GATTAATTTG ATTTCTTTAA TTCCTTATTG  2940
AATCTTATGT GATATGAATT TCTGGATTTA CAGAACATTA GCACATGTAC CTTGTGCCTC  3000
CCATTCAAGT GAAGTTATAA TTTACACTGA GGGTTTCAAA ATTCGACTAG AAGTGGAGAT  3060
ATATTATTTA TTTATGCACT GTACTGTATT TTTATATTGC TGTTTAAAAC TTTTAAGCTG  3120
TGCCTCACTT ATTTAAAGCAC AAAATGTTTT ACCTACTCCT TATTTACGAC ACAATAAAAT  3180
AACATCAATA GATTTTTAGG CTGAATTAAT TTGAAAGCAG CAATTTGCTG TTCTCAACCA  3240
TTCTTTCAAG GCTTTTCATT CGACACAATA AAATAACATC AATAG
```

Seq ID NO: 110 <u>Protein sequence</u>
Protein Accession #: NP_000484.2

```
1          11         21         31         41         51
|          |          |          |          |          |
MLPQIPFLLL VSLNLVHGVF YAERYQMPTG IKGPLPNTKT QFFIPYTIKS KGIAVRGEQG   60
TPGPPGPAGP RGHPGPSGPP GKPGYGSPGL QGEPGLPGPP GPSAVGKPGV PGLPGKPGER  120
GPYGPKGDVG PAGLPGPRGP PGPPGIPGPA GISVPGKPGQ QGPTGAPGPR GFPGEKGAPG  180
VPGMNGQKGE MGYGAPGRPG ERGLPGPQGP TGPSGPPGVG KRGENGVPGQ PGIKGDRGFP  240
GEMGPIGPPG PQGPPGSERGP EGIGKPGAAG APGQPGIPGT KGLPGAPGIA GPPGPPGFGK  300
PGLPGLKGER GPAGLPGGPG AKGEQGPAGL PGKPGLTGPP GNMGPQGPKG IPGSHGLPGP  360
KGETGPAGPA GYPGAKGERG SPGSDGKPGP GKPGLDGPK GNPGLPGPKG DPGVGGPPGL  420
PGFVGPAGAK GMPGHNGEAG PRGAPGIPGT RGPIGPPGIP GFPGSKGDPG SPGPPGPAGI  480
ATKGLNGPTG PPGPPGPRGH SGEPGLPGPP GPPGPPGQAV MPEGFIKAGQ RPSLSGTPLV  540
SANQGVTGMP VSAPTVILSK AYPAIGTPIP FDKILYNRQQ HYDPRTGIFT CQIPGIYYFS  600
YHVHVKGTHV WVGLYKNGTP VMYTYDEYTK GYLDQASGSA IIDLTENDQV WLQLPNAESN  660
GLYSSEYVHS SFSGFLVAPM
```

Seq ID NO: 111 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_000949
Coding sequence: 285..2153

```
        1         11         21         31         41         51
        |         |          |          |          |          |
GGAGGCTGAA ATCCCCAGAC GCCGGTTTTC TGGGCTGGGC TTTCTGCTTA CTCACTCCTT    60
CTCCCTCTTT CTGGATTTTA CCGACCGTTC GCGAAACAGC TTTCCACACA ATGGAGCTTC   120
ATGTCCTCGT GCAGGAAGTA CTCATCGACT GATGTGGCAG ACTTTGCTCC CTGACAAAAC   180
TAAAGAACTC TCCTATTCAT GGAGGCGAAC ACTGAGGATG CTTTCCACAT GAACCCTGAA   240
GTGAACTTCT GATACATTTC CTGCAGCAAG AGAAGGCAGC CAACATGAAG GAAAATGTGG   300
CATCTGCAAC CGTTTTCACT CTGCTACTTT TTCTCAACAC CTGCCTTCTG AATGGACAGT   360
TACCTCCTGG AAAACCTGAG ATCTTTAAAT GTCGTTCTCC CAATAAGGAA ACATTCACCT   420
GCTGGTGGAG GCCTGGGACA GATGGAGGAC TTCCTACCAA TTATTCACTG ACTTACCACA   480
GGGAAGGAGA GACACTCATG CATGAATGTC CAGACTACAT AACCGGTGGC CCCAACTCCT   540
GCCACTTTGG CAAGCAGTAC ACCTCCATGT GGAGGACATA CATCATGATG GTCAATGCCA   600
CTAACCAGAT GGGAAGCAGT TTCTCGGATG AACTTTATGT GGACGTGACT TACATAGTTC   660
AGCCAGACCC TCCTTTGGAG CTGGCTGTGG AAGTAAAACA GCCAGAAGAC AGAAAACCCT   720
ACCTGTGGAT TAAATGGTCT CCACCTACCC TGATTGACTT AAAAACTGGT TGGTTCACGC   780
TCCTGTATGA AATTCGATTA AAACCGGAGA AAGCAGCTGA GTGGGAGATC CATTTTGCTG   840
GGCAGCAAAC AGAGTTTAAG ATTCTCAGCC TACATCCAGG ACAGAAATAC CTTGTCCAGG   900
TTCGCTGCAA ACCAGACCAT GGATACTGGA GTGCATGGAG TCCAGCCGACC TTCATTCAGA   960
TACCTAGTGA CTTCACCATG AATGATACAA CCGTGTGGAT CTCTGTGGCT GTCCTTTCTG  1020
CTGTCATCTG TTTGATTATT GTCTGGGCAG TGGCTTTGAA GGGCTATAGC ATGGTGACCT  1080
GCATCTTTCC GCCAGTTCCT GGGCCAAAAA TAAAAGGATT TGATGCTCAT CTGTTGGAGA  1140
AGGGCAAGTC TGAAGAACTA CTGAGTGCCT TGGGATGCCA AGACTTTCCT CCCACTTCTG  1200
ACTATGAGGA CTTGCTGGTG GAGTATTTAG AAGTAGATGA TAGTGAGGAC CAGCATCTAA  1260
TGTCAGTCCA TTCAAAAGAA CACCCAAGTC AAGGTATGAA ACCCACATAC CTGGATCCTG  1320
ACACTGACTC AGGCCGGGGG AGCTGTGACA GCCCTTCCCT TTTGTCTGAA AAGTGTGAGG  1380
AACCCCAGGC CAATCCCTCC ACATTCTATG ATCCTGAGGT CATTGAGAAG CCAGAGAATC  1440
CTGAAACAAC CCACACCTGG GACCCCCAGT GCATAAGCAT GGAAGGCAAA ATCCCTATT   1500
TTCATGCTGG TGGATCCAAA TGTTCAACAT GGCCCTTACC ACAGCCCAGC CAGCACAACC  1560
CCAGATCCTC TTACCACAAT ATTACTGATG TGTGTGAGCT GGCTGTGGGC CCTGCAGGTG  1620
CACCGGCCAC TCTGTTGAAT GAAGCAGGTA AAGATGCTTT AAAATCCTCT CAAACCATTA  1680
AGTCTAGAGA AGAGGGAAAG GCAACCCAGC AGAGGGAGGT AGAAAGCTTC CATTCTGAGA  1740
CTGACCAGGA TACGCCCTGG CTGCTGCCCC AGGAGAAAAC CCCCTTTGGC TCCGCTAAAC  1800
CCTTGGATTA TGTGGAGATT CACAAGGTCA ACAAAGATGG TGCATTATCA TTGCTACCAA  1860
AACAGAGAGA GAACAGCGGC AAGCCCAAGA AGCCCGGGAC TCCTGAGAAC AATAAGGAGT  1920
ATGCCAAGGT GTCCGGGGTC ATGGATAACA ACATCCTGGT GTTGGTGCCA GATCCACATG  1980
CTAAAAACGT GGCTTGCTTT GAAGAATCAG CCAAAGAGGC CCCACCATCA CTTGAACAGA  2040
ATCAAGCTGA GAAAGCCCTG GCCAACTTCA CTGCAACATC AAGCAAGTGC AGGCTCCAGC  2100
TGGGTGGTTT GGATTACCTG GATCCCGCAT GTTTTACACA CTCCTTTCAC TGATAGCTTG  2160
ACTAATGGAA TGATTGGTTA AAATGTGATT TTTCTTCAGG TAACACTACA GAGTACGTGA  2220
AATGCTCAAG AATGTAGTCA GACTGACACT ACTAAAGCTC CCAGCTCCTT TCATGCTCCA  2280
TTTTAACCA CTTGCCTCTT TCTCCAGCAG CTGATTCCAG AACAAATCAT TATGTTTCCT  2340
AACTGTGATT TGTAGATTTA CTTTTTGCTG TTAGTTATAA AACTATGTGT TCAATGAAAT  2400
AAAAGCACAC TGCTTAGTAT TCTTGAGGGA CAATGCCAAT AGGTATATCC TCTGGAAAAG  2460
GCTTTCATGA TTTGGCATGG GACAGACGGA AATGAAATTG TCAAAATTGT TTACCATAGA  2520
AAGATGACAA AAGAAAATTT TCCACATAGG AAAATGCCAT GAAAATTGCT TTTGAAAAAC  2580
AACTGCATAA CCTTTACACT CCTCGTCCAT TTTATTAGGA TTACCCAAAT ATAACCATTT  2640
AAAGAAAGAA TGCATTCCAG AACAAATTGT TTACATAAGT TCCTATACCT TACTGACACA  2700
TTGCTGATAT GCAAGTAAGA AAT
```

Seq ID NO: 112 <u>Protein sequence</u>
Protein Accession #: NP_000940

```
        1          11          21         31          41          51
        |          |           |          |           |           |
MKRTVASATV FTLLLFLNTC LLNGQLPPGK PEIFKCRSPN KETFTCWWRP GTDGGLPTNY    60
SLTYHREGET LMHECPDYIT GGPNSCHFGK QYTSMWRTYI MMVNATNQMG SSFSDELYVD   120
VTYIVQPDPP LELAVEVKQP EDRKPYLWIK WSPPTLIDLK TGWFTLLYEI RLKPEKAAEW   180
EIHFAGQQTE FKILSLHPGQ KYLVQVRCKP DHGYNSAWSP ATFIQIPSDF TMNDTTVWIS   240
VAVLSAVICL IIVWAVALKG YSMVTCIFPP VPGPKIKGFD AHLLEKGKSE ELLSALGCQD   300
FPPTSDYEDL LVEYLEVDDS EDQHLMSVHS KEHPSQGMKP TYLDPDTDSG RGSCDSPSLL   360
SEKCEEPQAN PSTFYDPEVI EKPENPETTH TWDPQCISME GKIPYPHAGG SKCSTWPLPQ   420
PSQHNPRSSY HNITDVCELA VGPAGAPATL LNEAGKDALK SSQTIKSREE GKATQQREVE   480
SFHSETDQDT PWLLPQEKTP FGSAKPLDYV EIHKVNKDGA LSLLPKQREN SGKPKKPGTP   540
ENNKEYAKVS GVMDNNILVL VPDPHAKNVA CFEESAKEAP PSLEQNQAEK ALANFTATSS   600
KCRLQLGGLD YLDPACFTHS FH
```

Seq ID NO: 113 <u>DNA sequence</u>
Nucleic Acid Accession #: XM_062811
Coding sequence: 1..888

```
        1          11          21         31          41          51
        |          |           |          |           |           |
ATGTGGGGCG CTCGCCGCTC GTCCGTCTCC TCATCCTGGA ACGCCGCTTC GCTCCTGCAG    60
CTGCTGCTGG CTGCGCTGCT GGCGGCGGGG GCGAGGGCCA GCGGCGAGTA CTGCCACGGC   120
TGGCTGGACG CGCAGGGCGT CTGGCGCATC GGCTTCCAGT GTCCCGAGCG CTTCGACGGC   180
GGCGACGCCA CCATCTGCTG CGGCAGCTGC GCGGTGCGGCT ACTGCTGCTC CAGCGCGCAG   240
GCGCGCCTGG ACCAGGGCGG CTGCGACAAT GACCGGCCAGC AGGGCGCTGG CGAGCCTGGC   300
CGGGCGGACA AAGACGGCCC CGACGGCTCG GCAGTGCCCA TCTACGTGCC GTTCCTCATT   360
GTTGGCTCCG TGTTTGTCGC CTTTATCATC TTGGGGTCCC TGGTGGCAGC CTGTTGCTGC   420
```

```
              AGATGTCTCC GGCCTAAGCA GGATCCCCAG CAGAGCCGAG CCCCAGGGGG TAACCGCTTG   480
              ATGGAGACCA TCCCCATGAT CCCCAGTGCC AGCACCTCCC GGGGGTCGTC CTCACGCCAG   540
              TCCAGCACAG CTGCCAGTTC CAGCTCCAGC GCCAACTCAG GGGCCCGGGC GCCCCCAACA   600
    5         AGGTCACAGA CCAACTGTTG CTTGCCGGAA GGGACCATGA ACAACGTGTA TGTCAACATG   660
              CCCACGAATT TCTCTGTGCT GAACTGTCAG CAGGCCACCG AGATTGTGCC ACATCAAGGG   720
              CAGTATCTGC ATCCCCCATA CGTGGGGTAC ACGGTGCAGC ACGACTCTGT GCCCATGACA   780
              GCTGTGCCAC CTTTCATGGA CGGCCTGCAG CCTGGCTACA GGCAGATTCA GTCCCCCTTC   840
              CCTCACACCA ACAGTGAACA GAAGATGTAC CCAGCGGTGA CTGTATAA
```

Seq ID NO: 114 Protein sequence
Protein Accession #: XP_062811

```
    1          11         21         31         41         51
    |          |          |          |          |          |
    MMGARRSSVS SSWNAASLLQ LLLAALLAAG ARASGEYCHG WLDAQGVWRI GFQCPERFDG   60
    GDATICCGSC ALRYCCSSAE ARLDQGGCDN DRQQGAGEPG RADKDGPDGS AVPIYVPFLI   120
    VGSVFVAFII LGSLVAACCC RCLRPKQDPQ QSRAPGGNRL METIPMIPSA STSRGSSSRQ   180
    SSTAASSSSS ANSGARAPPT RSQTNCCLPE GTMNNVYVNM PTNFSVLNCQ QATQIVPHQG   240
    QYLHPPYVGY TVQHDSVPMT AVPPFMDGLQ PGYRQIQSPF PHTNSEQKMY PAVTV
```

Seq ID NO: 115 DNA sequence
Nucleic Acid Accession #: NM_013257
Coding sequence: 223..1512

```
    1          11         21         31         41         51
    |          |          |          |          |          |
    GGTGTGCTCT TGAGGGATTA AATGCAAAGA GATCACACCA TGGACTACAA GGAAAGCTGC   60
    CCAAGTGTAA GCATTCCCAG CTCCGATGAA CACAGAGAGA AAAAGAAGAG GTTTACTGTT   120
    TATAAAGTTC TGGTTTCAGT GGGAAGAAGT GAATGGTTTG TCTTCAGGAG ATATGCAGAG   180
    TTTGATAAAC TTTATAACAC TTTAAAAAAA CAGTTTCCTG CTATGGCCCT GAAGATTCCT   240
    GCCAAGAGAA TATTTGGTGA TAATTTTGAT CCAGATTTTA TTAAACAAAG ACGAGCAGGA   300
    CTAAACGAAT TCATTCAGAA CCTAGTTAGG TATCCAGAAC TTTATAACCA TCCAGATGTC   360
    AGAGCATTCC TTCAAATGGA CAGTCCAAAA CACCAGTCAG ATCCATCTGA AGATGAGGAT   420
    GAAAGAAGTT CTCAGAAGCT ACACTCTACC TCACAGAACA TCAACCTGGG ACCGTCTGGA   480
    AATCCTCATG CCAAACCAAC TGACTTTGAT TTCTTAAAAG TTATTGGAAA AGGCAGCTTT   540
    GGCAAGGTTC TTCTTGCAAA ACGGAAACTG GATGGAAAAT TTTATGCTGT CAAAGTGTTA   600
    CAGAAAAAAA TAGTTCTCAA CAGAAAGAG CAAAAACATA TTATGGCTGA ACGTAATGTG   660
    CTCTTGAAAA ATGTGAAACA TCCGTTTTTG GTTGGATTGC ATTATTCCTT CCAAACAACT   720
    GAAAAGCTTT ATTTTGTTCT GGATTTTGTT AATGGCAGGG AGCTTTTTGA CCACTTACAA   780
    AGAGAACGGT CCTTTCCTGA GCACAGAGCT AGGTTTTACG CTGCTGAAAT TGCTAGTGCA   840
    TTGGGTTACT TACATTCCAT CAAAATAGTA TACAGAGACT TGAAACCAGA AAATATTCTT   900
    TTGGATTCAG TAGGACATGT TGTCTTAACA GATTTTGGGC TTTGTAAAGA AGGAATTGCT   960
    ATTTCTGACA CCACTACCAC ATTTGTGGG ACACCAGAGT ATCTTGCACC TGAAGTAATT   1020
    AGAAAACAGC CCTATGACAA TACTGTAGAT TGGTGGTGCC TTGGGGCTGT TCTGTATGAA   1080
    ATGCTGTATG GATTGCCTCC TTTTTATTGC CGAGATGTTG CTGAAATGTA TGACAATATC   1140
    CTTCACAAAC CCCTAAGTTT GAGGCCAGGA GTGAGTCTTA CAGCCTGGTC CATTCTGGAA   1200
    GAACTCCTAG AAAAAGACAG GCAAAATCGA CTTGGTGCCA AGGAAGACTT TCTTGAAATT   1260
    CAGAATCATC CTTTTTTTGA ATCACTCAGC TGGGCTGACC TTGTACAAAA GAAGATTCCA   1320
    CCACCATTTA ATCCTAATGT GGCTGGAGCA GATGATATCA GAAACTTTGA CACAGCATTT   1380
    ACAGAAGAAA CAGTTCCATA TTCTGTGTGT GTATCTTCTG ACTATTCTAT AGTGAATGCC   1440
    AGTGTATTGG AGGCAGATGA TGCATTCGTT GGTTTCTCTT ATGCACCTCC TTCAGAAGAC   1500
    TTATTTTTGT GAGCAGTTTG CCATTCAGAA ACCATTGAGC AAAATAAGTC TATAGATGGG   1560
    ACTGAAACTT CTATTTGTGT GAATATAATTC AAATATGTAT AACTAGTGCC TCATTTTTAT   1620
    ATGTAATGAT GAAAACTATG AAAAAAATGTA TTTTCTTCTA TGTGCAAGAA AAATAGGGCA   1680
    TTTCAAAGAG CTGTTTTGAT TAAAATTTAT ATTCTTGTTT AATAAGCTTA TTTTTAAACA   1740
    ATTTAAAAGC TATTATTCTT AGCATTAACC TATTTTTAAA GAAACCTTTT TTGCTATTGA   1800
    CTGTTTTTTC CCTCTAAGTT TACACTAACA TCTACCCAAG ATAGACTGTT TTTTAACAGT   1860
    CAATTTCAGT TCAGCTAACA TATATTAATA CCTTTGTAAC TCTTTGCTAT GGCTTTTGTT   1920
    ATCACACCAA AACTATGCAA TTGGTACATG GTTGTTTAAG AAGAAACCGT ATTTTTTCCAT   1980
    GATAAATCAC TGTTTGAAAT ATTTGGTTCA TGGTATGATC GAAATGTAAA AGCATAATTA   2040
    ACACATTGGC TGCTAGTTAA CAATTGGAAT AACTTTATTC TGCAGATCAT TTAAGAAGTA   2100
    ACAGGCCGGG CGCGGTGGCT CACGCCTGTA ATCCCAGCAC TTTGGGAGGC TGAGGCGGGC   2160
    AGATCACCTG AGGTCAGGAG TTGGAGACCA GCCTGACCAA CATGGACAAA CCCCGTCTCT   2220
    ACTAAAAATA CAAAATTGGC AGGGTGTGGT GGCACATGCC TATAATCCCA GCTACTTGGG   2280
    AGGCTAAGGC AGGAGAATCG CTTGAACCCG GGAGGCGGAG GTTGCAGTGA GCCGAGATCG   2340
    CACCATTGCA CTCCTGCCTG GGCAACAAGA GTGAAACTCC ATCTCCAAAA A
```

Seq ID NO: 116 Protein sequence
Protein Accession #: NP_037389

```
    1          11         21         31         41         51
    |          |          |          |          |          |
    MALKIPARRI FGDNFDPDFI KQRRAGLNEF IQNLVRYPEL YNHPDVRAFL QMDSPKHQSD   60
    PSEDEDERSS QKLHSTSQNI NLGPSGNPHA KPTDFDFLKV IGKGSFGKVL LAKRKLDGKF   120
    YAVKVLQKKI VLNRKEQKHI MAERNVLLKN VKHPFLVGLH YSFQTTEKLY FVLDFVNGGE   180
    LFFHLQRERS FPEHRARFYA AEIASALGYL HSIKIVYRDL KPENILLDSV GHVVLTDFGL   240
    CKEGIAISDT TTTFCGTPEY LAPEVIRKQP YDNTVDWWCL GAVLYEMLYG LPPFYCRDVA   300
    EMYDNILHKP LSLRPGVSLT AWSILEELLE KDRQNRLGAK EDFLEIQNHP FFESLSWADL   360
    VQKKIPPPFN PNVAGPDDIR NFDTAFTEET VPYSVCVSSD YSIVNASVLE ADDAFVGFSY   420
    APPSEDLFL
```

Seq ID NO: 117 DNA sequence
Nucleic Acid Accession #: NM_004004.1

Coding sequence: 1..681

```
        1           11          21          31          41          51
        |           |           |           |           |           |
 5   ATGGATTGGG GCACGCTGCA GACGATCCTG GGGGGTGTGA ACAAACACTC CACCAGCATT    60
     GGAAAGATCT GGCTCACCGT CCTCTTCATT TTTCGCATTA TGATCCTCGT TGTGGCTGCA   120
     AAGGAGGTGT GGGGAGATGA GCAGGCCGAC TTTGTCTGCA ACACCCTGCA GCCAGGCTGC   180
     AAGAACGTGT GCTACGATCA CTACTTCCCC ATCTCCCACA TCCGGCTATG GGCCCTGCAG   240
     CTGATCTTCG TGTCCAGCCC AGCGCTCCTA GTGGCCATGC ACGTGGCCTA CCGGAGACAT   300
10   GAGAAGAAGA GGAAGTTCAT CAAGGGGGAG ATAAAGAGTG AATTTAAGGA CATCGAGGAG   360
     ATCAAAACCC AGAAGGTCCG CATCGAAGGC TCCCTGTGGT GGACCTACAC AAGCAGCATC   420
     TTCTTCCGGG TCATCTTCGA AGCCGCCTTC ATGTACGTCT CTATGTCAT GTACGACGGC    480
     TTCTCCATGC AGCGGCTGGT GAAGTGCAAC GCCCTGGCCTT GTCCCAACAC TGTGGACTGC   540
     TTTGTGTCCC GGCCCACGGA GAAGACTGTC TTCACAGTGT TCATGATTGC AGTGTCTGGA   600
15   ATTTGCATCC TGCTCGAATGT CACTGAATTG TGTTATTTGC TAATTAGATA TTGTTCTGGG   660
     AAGTCAAAAA AGCCAGTTTA A
```

Seq ID NO: 118 <u>Protein sequence</u>
Protein Accession #: NP_003995.1

```
        1           11          21          31          41          51
        |           |           |           |           |           |
     MDWGTLQTIL GGVNKHSTSI GKIWLTVLFI FRIMILVVAA KEVWGDEQAD FVCNTLQPGC    60
     KNVCYDHYFP ISHIRLWALQ LIFVSSPALL VAMHVAYRRH EKRKFIKGE IKSEFKDIEE    120
25   IKTQKVRIEG SLWWTYTSSI FFRVIFEAAF MYVFYVMYDG FSMQRLVKCN AWPCPNTVDC   180
     FVSRPTEKTV FTVFMIAVSG ICILLNVTEL CYLLIRYCSG KSKKPV
```

Seq ID NO: 119 <u>DNA sequence</u>
Nucleic Acid Accession #: XM_061091.1
Coding sequence: 1..2481

```
        1           11          21          31          41          51
        |           |           |           |           |           |
35   ATGCCAAATA CTTCAGGAAC AACCAGGATT GAAATTTGGC TTCTCCAAGA GCCGCCCGGG    60
     CACCGAGCGC TGGTCGCCGC TCTCCTTCCG GTGAGTCCCA GCCCCGAGTT GGCTCTGGCG   120
     CCCGGGTACC CGCCAGTGCC GGCTGCCGAT GACCGATTCA CGCTCCCGAT GATTGGAGGT   180
     CAGATGCATG GTGAGAAGGT AGATCTCTGG AGCCTTGGTG TTCTTTGCTA TGAATTTTTA   240
     GTTGGGAAGC CTCCTTTTGA GGCAAACGAA GTCCATGTAA GCAAAGAAAC CATCGGGAAG   300
40   ATTTCAGCTG CCAGCAAAAT GATGTGGTGC TCGGCTGCAG TGGACATCAT GTTTCTGTTA   360
     GATGGGTCTA ACAGCGTCGG GAAAGGGAGC TTTGAAAGGT CCAAGCACTT TGCCATCACA   420
     GTCTGTGACG GTCTGGACAT CAGCCCCGAG AGGGTCAGAG TGGGAGCATT CCAGTTCAGT   480
     TCCACTCCTC ATCTGGAATT CCCCTTGGAT TCATTTTCAA CCCAACAGGA AGTGAAGGCA   540
     AGAATCAAGA GGATGGTTTT CAAAGGAGGG CGCACGGAGA CGGAACTTGC TCTGAAATAC   600
45   CTTCTGCACA GAGGGTTGCC TGGAGGCAGA AATGCTTCTG TGCCCAGAT CCTCATCATC     660
     GTCACTGATG GGAAGTCCCA GGGGGGATGTG GCACTGCCAT CCAAGCAGCT GAAGGAAAGG   720
     GGTGTCACTG TGTTTGCTGT GGGGGTCAGG TTTCCCAGGT GGGAGGAGCT GCATGCACTG   780
     GCCAGCGAGC CTAGAGGGCA GCACGTGCTG TTGGCTGAGC AGGTGGAGGA TGCCACCAAC   840
     GGCCTCTTCA GCACCCTCAG CAGCTCGGCC ATCTGCTCCA GCGCCACGCC AGCTGGGAGC   900
50   CCCGAGCTTG TCTTCATGGA GCGGTTAATG GGCATCTCTC TGATAGGCCC CTGTGACTCG   960
     CAGCCCTGCC AGAATGGAGG CACATGTGTT CCAGAAGGAC TGGACGGCTA CCAGTGCCTC  1020
     TGCCCGCTGG CCTTTGGAGG GGAGGCTAAC TGTGCCCTGA AGCTGAGCCT GGAATGCAGG  1080
     GTCGACCTCC TCTTCCTGCT GGACAGCTCT GCGGGCACCA CTCTGGACGG CTTCCTGCGG  1140
     GCCAAAGTCT TCGTGAAGCG GTTTGTGCGG GCCGTGCTGA GCGAGGACTC TCGGGCCCGA  1200
55   GTGGGTGTGG CCACATACAG CAGGGAGCTG CTGGTGGCGG TGCCTGTGGG GGAGTACCAG  1260
     GATGTGCCTG ACCTGGTCTG GAGCCTCGAT GGCATTCCCT TCCGTGGTGG CCCCACCCTG  1320
     ACGGGCAGTG CCTTGCGGCA GGCGGCAGAG CGTGGCTTCG GGAGCGCCAC CAGGACAGGC  1380
     CAGGACCGGC CACGTAGAGT GGTGGTTTTG CTCACTGAGT CACACTCCGA GGATGAGGTT  1440
     GGGGGCCCAG CGCGTCACGC AAGGGCGCGA GAGCTGCTCC TGCTGGGGTGT AGGCAGTGAG  1500
60   GCCGTGCGGG CAGAGCTGGA GGAGATCACA GGCAGCCCAA AGCATGTGAT GGTCTACTCG  1560
     GATCCTCAGG ATCTGTTCAA CCAAATCCCT GAGCTGCAGG GGAAGCTGTG CAGCCGGCAG  1620
     CGGCTAGGGT GCGGGACACA AGCCCTGGAC CTCGTCTTCA TGTTGGACAC CTCTGCCTCA  1680
     GTAGGGCCCG AGAATTTTGC TCAGATGCAG AGCTTTGTGA GAAGCTGTGC CCTCCAGTTT  1740
     GAGGTGAACC CTGACGTGAC ACAGGTCGGC CTGGTGGTGT ATGGCAGCCA GGTGCAGACT  1800
65   GCCTTCGGGC TGGACACCAA ACCCACCCGG GCTGCGATGC TGCGGGCCAT TAGCCAGGCC  1860
     CCCTACCTAG GTGGGGTGGG CTCAGCCGGC ACCGCCCTGC TGCACATCTA TGACAAAGTG  1920
     ATGACCGTCC AGAGGGGTGC CCGGCCTGGT GTCCCCAAAG CTGTGGTGGT GCTCACAGGC  1980
     GGGAGAGGCG CAGAGGATGC CGACGTTCCT GCCCAGAAGC TGAGGAACAA TGGCATCTCT  2040
     GTCTTGGTCG TGGGCGTGGG GCCTGTCCTA AGTGAGGGTC TGCGGAGGCT TGCAGGTCCC  2100
70   CGGGATTCCC TGATCCACGT GGCAGCTTAC GCCGACCTGC GGTACCACCA GGACGTGCTC  2160
     ATTGAGTGGC TGTGTGGAGA AGCCAAGCAG CCAGTCAACC TCTGCAAACC TGCTGGGAGC  2220
     ATGAATGAGG GCAGCTGCGT CCTGCAGAAT GGGAGCTACC GCTGCAAGTG TCGGGATGGC  2280
     TGGGAGGGGCC CCCACTGCGA GAACCGTGAG TGGAGCTCTT GCTCTGTATG TGTGAGCCAG  2340
     GGATGGATTC TTGAGACGCC CCTGAGGCAC ATGGCTCCCG TGCAGGAGGG CAGCAGCCGT  2400
75   ACCCCTCCCA GCAACTACAG AGAAGGCCTG GGCACTGAAA TGGTGCCTAC CTTCTGGAAT  2460
     GTCTGTGCCC CAGGTCCTTA G
```

Seq ID NO: 120 <u>Protein sequence</u>
Protein Accession #: XP_061091.1

```
        1           11          21          31          41          51
        |           |           |           |           |           |
     MPNTSGTTRI EIWLLQEPPG HRALVAALLP VSPSPELALA PGYPPVPAAD DRFTLPMIGG    60
     QMHGEKVDLW SLGVLCYEFL VGKPPFEANE VHVSKETIGK ISAASKMMWC SAAVDIMFLL   120
```

```
    DGSNSVGKGS  FERSKHFAIT  VCDGLDISPE  RVRVGAFQFS  STPHLEFPLD  SPSTQQEVKA   180
    RIKRMVFKGG  RTETELALKY  LLHRGLPGGR  NASVPQILII  VTDGKSQGDV  ALPSKQLKER   240
    GVTVFAVGVR  PPRWEELHAL  ASEPRGQHVL  LAEQVEDATN  GLFSTLSSSA  ICSSATPAGS   300
    PELVPMERLM  GISLIGPCDS  QPCQNGGTCV  PEGLDYQCL   CPLAFGGEAN  CALKLSLECR   360
    VDLLFLLDSS  AGTTLDGFLR  AKVFVKRFVR  AVLSEDSRAR  VGVATYSREL  LVAVPVGEYQ   420
    DVPDLVWSLD  GIPFRGGPTL  TGSALRQAAE  RGFGSATRTG  QDRPRRVVVL  LTESHSEDEV   480
    AGPARHARAR  ELLLLGVGSE  AVRAELEEIT  GSPKHVMVYS  DPQDLFNQIP  ELQGKLCSRQ   540
    RPGCRTQALD  LVFMLDTSAS  VGPENFAQMQ  SFVRSCALQF  EVNPDVTQVG  LVVYGSQVQT   600
    AFGLDTKPTR  AAMLRAISQA  PYLGGVGSAG  TALLHIYDKV  MTVQRGARPG  VPKAVVVLTG   660
    GRGAEDAAVP  AQKLRNNGIS  VLVVGVGPVL  SEGLRRLAGP  RDSLIHVAAY  ADLRYHQDVL   720
    IEWLCGEAKQ  PVNLCKPSPC  MNEGSCVLQN  GSYRCKCRDG  WEGPHCENRE  WSSCSVCVSQ   780
    GWILETPLRH  MAPVQEGSSR  TPPSNYREGL  GTEMVPTFWN  VCAPGP
```

Seq ID NO: 121 <u>DNA sequence</u>
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..2424

```
    1          11         21         31         41         51
    |          |          |          |          |          |
    ATGCCCCTT  TCCTGTTGCT  GGAGGCCGTC  TGTGTTTTCC  TGTTTTCCAG  AGTGCCCCCA   60
    TCTCTCCCTC  TCCAGGAAGT  CCATGTAAGC  AAAGAAACCA  TCGGGAAGAT  TTCAGCTGCC  120
    AGCAAATGA   TGTGGTGCTC  GGCTGCAGTG  GACATCATGT  TTCTGTTAGA  TGGGTCTAAC  180
    AGCGTCGGGA  AAGGGAGCTT  TGAAAGGTCC  AAGCACTTTG  CCATCACAGT  CTGTGACGGT  240
    CTGGACATCA  GCCCCGAGAG  GGTCAGAGTG  GGAGCATTCC  AGTTCAGTTC  CACTCCTCAT  300
    CTGGAATTCC  CCTTGGATTC  ATTTTCAACC  CAACAGGAAG  TGAAGGCAAG  AATCAAGAGG  360
    ATGGTTTTCA  AAGGAGGGCG  CACGGAGACG  GAACTTGCTC  TGAAATACCT  TCTGCACAGA  420
    GGGTTGCCTG  GAGGCAGAAA  TGCTTCTGTG  CCCAGATCC   TCATCATCGT  CACTGATGGG  480
    AAGTCCCAGG  GGGATGTGGC  ACTGCCATCC  AAGCAGCTGA  AGGAAAGGGG  TGTCACTGTG  540
    TTTGCTGTGG  GGGTCAGGTT  TCCCAGGTGG  GAGGAGCTGC  ATGCACTGGC  CAGCGAGCCT  600
    AGAGGGCAGC  ACGTGCTGTT  GGCTGAGCAG  GTGGAGGATG  CCACCAACGG  CCTCTTCAGC  660
    ACCCTCAGCA  GCTCGGCCAT  CTGCTCCAGC  GCCACGCCAG  ACTGCAGGGT  CGAGGCTCAC  720
    CCCTGTGAGC  ACAGGACGCT  GGAGATGGTC  CGGGAGTTCG  CTGGCAATGC  CCCATGCTGG  780
    AGAGGATCGC  GGCGGACCCT  TGCCGTGCTG  GCTGCACACT  GTCCCTTCTA  CAGCTGGAAG  840
    AGAGTGTTCC  TAACCCACCC  TGCCACCTGC  TACAGGACCA  CCTGCCCAGG  CCCCTGTGAC  900
    TCGCAGCCCT  GCCAGAATGG  AGGCACATGT  GTTCCAGAAG  GACTGGACGG  CTACCAGTGC  960
    CTCTGCCCGC  TGGCCTTTGG  AGGGGAGGCT  AACTGTGCCC  TGAAGCTGAG  CCTGGAATGC  1020
    AGGGTCGACC  TCCTCTTCCT  GCTGGACAGC  TCTGCGGGCA  CCACTCTGGA  CGGCTTCCTG  1080
    CTGGCCAAAG  TCTTCGTGAA  GCGGTTTGTG  CGGGCCGTGC  TGAGCGAGGA  CTCTCGGGCC  1140
    CGAGTGGGTG  TGGCCACATA  CAGCAGGGAG  CTGCTGGTGG  CGGTGCCTGT  GGGGGAGTAC  1200
    CAGGATGTGC  CTGACCTGGT  CTGGAGCCTC  GATGGCATTC  CCTTCCGTGG  TGGCCCCACC  1260
    CTGACGGGCA  GTGCCTTGCG  GCAGGCGGCA  GAGCGGTGGC  TCGGGAGCGC  CACCAGGACA  1320
    GGCCAGGACC  GGCCACGTAG  AGTGGTGGTT  TTGCTCACTG  AGTCACACTC  CGAGGATGAG  1380
    GTTGCGGGCC  CAGCGCGTCA  CGCAAGGGCG  CGAGAGCTGC  TCCTGCTGGG  TGTAGGCAGT  1440
    GAGGCCGTGC  GGGCAGAGCT  GGAGGAGATC  ACAGGCAGCC  CAAAGCATGT  GATGGTCTAC  1500
    TCGGATCCTC  AGGATCTGTT  CAACCAAATC  CCTGAGCTGC  AGGGGGAAGCT  GTGCAGCCGG  1560
    CAGCGGCCAG  GGTGCCGGAC  ACAAGCCCTG  GACCTGGTCT  TCATGTTGGA  CACCTCTGCC  1620
    TCAGTAGGGC  CCGAGAATTT  TGCTCAGATG  CAGAGCTTTG  TGAGAAGCTG  TGCCCTCCAG  1680
    TTTGAGGTGA  ACCCTGACGT  GACACAGGTC  GGCCTGGTGG  TGTATGGCAG  CCAGGTGCAG  1740
    ACTGCCTTCG  GGCTGGACAC  CAAACCCACC  CGGGCTGCGA  TGCTGCGGGC  CATTAGCCAG  1800
    GCCCCCTACC  TAGGTGGGGT  GGGCTCAGCC  GGCACGGCCC  TGCTGCACAT  CTATGACAAA  1860
    GTGATGACCG  TCCAGAGGGG  TGCCCGGCCT  GGTGTCCCCA  AAGCTGTGGT  GGTGCTCACA  1920
    GGCGGGGAGG  GCGCAGAGGA  TGCAGCCGTT  CCTGCCCAGA  AGCTGAGGAA  CAATGGCATC  1980
    TCTGTCTTGG  TCGTGGGCGT  GGGGCCTGTC  CTAAGTGAGG  GTCTGCGGAG  GCTTGCAGGT  2040
    CCCGGGATT   CCCTGATCCA  CGTGGCAGCT  TACGCCGACC  TGCGGTACCA  CCAGGACGTG  2100
    CTCATTGAGT  GGCTGTGTGG  AGAAGCCAAG  CAGCCAGTCA  ACCTCTGCAA  ACCCAGCCCG  2160
    TGCATGAATG  AGGGCAGCTG  CGTCCTGCAG  AATGGGAGCT  ACCGCTGCAA  GTGTCGGGAT  2220
    GGCTGGGAGG  GCCCCCACTG  CGAGAACCGT  GAGTGGAGCT  CTTGCTCTGT  ATGTGTGAGC  2280
    CAGGGATGGA  TTCTTGAGAC  GCCCCTGAGG  CACATGGCTC  CCGTGCAGGA  GGGCAGCAGC  2340
    CGTACCCCTC  CCAGCAACTA  CAGAGAAGGC  CTGGGCACTG  AAATGGTGCC  TACCTTCTGG  2400
    AATGTCTGTG  CCCCAGGTCC  TTAG
```

Seq ID NO: 122 <u>Protein sequence</u>
Protein Accession #: Eos sequence

```
    1          11         21         31         41         51
    |          |          |          |          |          |
    MPPFLLLEAV  CVFLFSRVPP  SLPLQEVHVS  KETIGKISAA  SKMMWCSAAV  DIMFLLDGSN   60
    SVGKGSFERS  KHFAITVCDG  LDISPERVRV  GAFQFSSTPH  LEFPLDSPST  QQEVKARIKR  120
    MVFKGGRTET  ELALKYLLHR  GLPGGRNASV  PQILIIVTDG  KSQGDVALPS  KQLKERGVTV  180
    FAVGVRPPRN  EELHALASEP  RGQHVLLAEQ  VEDATNGLFS  TLSSSAICSS  ATPDCRVEAH  240
    PCEHRTLEMV  REFAGNAPCW  RGSRRTLAVL  AAHCPFYSWK  RVFLTHPATC  YRTTCPGPCD  300
    SQPCQNGGTC  VPEGLDGYQC  LCPLAFGGEA  NCALKLSLEC  RVDLLFLLDS  SAGTTLDGFL  360
    RAKVFVKRFV  RAVLSEDSRA  RVGVATYSRE  LLVAVPVGEY  QDVPDLVWSL  DGIPFRGGPT  420
    LTGSALRQAA  ERGFGSATRT  GQDRPRRVVV  LLTESHSEDE  VAGPARHARA  RELLLLGVGS  480
    EAVRAELEEI  TGSPKHVMVY  SDPQDLFNQI  PELQGKLCSR  QRPGCRTQAL  DLVFMLDTSA  540
    SVGPENFAQM  QSFVRSCALQ  FEVNPDVTQV  GLVVYGSQVQ  TAFGLDTKPT  RAAMLRAISQ  600
    APYLGGVGSA  GTALLHIYDK  VMTVQRGARP  GVPKAVVVLT  GGRGAEDAAV  PAQKLRNNGI  660
    SVLVVGVGPV  LSEGLRRLAG  PRDSLIHVAA  YADLRYHQDV  LIEWLCGEAK  QPVNLCKPSP  720
    CMNEGSCVLQ  NGSYRCKCRD  GWEGPHCENR  EWSSCSVCVS  QGWILETPLR  HMAPVQEGSS  780
    RTPPSNYREG  LGTEMVPTFW  NVCAPGP
```

Seq ID NO: 123 <u>DNA sequence</u>
Nucleic Acid Accession #: Eos sequence
Coding sequence: 89..2356

```
        1           11          21          31          41          51
        |           |           |           |           |           |
5       GCCCCCTGGC  CCGAGCCGCG  CCCGGGTCTG  TGAGTAGAGC  CGCCCGGGCA  CCGAGCGCTG     60
        GTCGCCGCTC  TCCTTCCGTT  ATATCAACAT  AGCCCCCTTTC CTGTTGCTGG  AAGCCGTCTG    120
        TGTTTTCCTG  TTTTCCAGAG  TGCCCCCATC  TCTCCCTCTC  CAGGAAGTCC  ATGTAAGCAA    180
        AGAAACCATC  GGGAAGATTT  CAGCTGCCAG  CAAAATGATG  TGGTGCTCGG  CTGCAGTGGA    240
        CATCATGTTT  CTGTTAGATG  GGTCTAACAG  CGTCGGGAAA  GGGAGCTTTG  AAAGGTCCAA    300
10      GCACTTTGCC  ATCACAGTCT  GTGACGGTCT  GGACATCAGC  CCCGAGAGGG  TCAGAGTGGG    360
        AGCATTCCAG  TTCAGTTCCA  CTCCTCATCT  GGAATTCCCC  TTGGATTCAT  TTTCAACCCA    420
        ACAGGAAGTG  AAGGCAAGAA  TCAAGAGGAT  GGTTTTCAAA  GGAGGGCGCA  CGGAGACGGA    480
        ACTTGCTCTG  AAATACCTTC  TGCACAGAGG  GTTGCCTGGA  GGCAGAAATG  CTTCTGTGCC    540
        CCAGATCCTC  ATCATCGTCA  CTGATGGGAA  GTCCCAGGGG  GATGTGGCAC  TGCCATCCAA    600
15      GCAGCTGAAG  GAAAGGGGTG  TCACTGTGTT  TGCTGTGGGG  GTCAGGTTTC  CCAGGTGGGA    660
        GGAGCTGCAT  GCACTGGCCA  GCGAGCCTAG  AGGGCAGCAG  GTGCTGTTGG  CTGAGCAGGT    720
        GGAGGATGCC  ACCAACGGCC  TCTTCAGCAC  CCTCAGCAGC  TCGGCCATCT  GCTCCAGCGC    780
        CACGCCAGAC  TGCAGGGTCG  AGGCTCACCC  CTGTGAGCAC  AGGACGCTGG  AGATGGTCCG    840
        GGAGTTCGCT  GGCAATGCCC  CATGCTGGAG  AGGATCGGCG  CGGACCCTTG  CGGTGCTGGC    900
20      TGCACACTGT  CCCTTCTACA  GCTGGAAGAG  AGTGTTCCTA  ACCCACCCTG  CCACCTGCTA    960
        CAGGACCACC  TGCCCAGGCC  CCTGTGACTC  GCAGCCCTGC  CAGAATGGAG  GCACATGTGT   1020
        TCCAGAAGGA  CTGGACGGCT  ACCAGTGCCT  CTGCCCGCTG  GCCTTTGGAG  GGGAGGCTAA   1080
        CTGTGCCCTG  AAGCTGAGCC  TGGAATGCAG  GGTCGACCTC  CTCTTCCTGC  TGGACAGCTC   1140
        TGCGGGCACC  ACTCTGGACG  GCTTCCTGCG  GGCCAAAGTC  TTCGTGAAGC  GGTTTGTGCG   1200
25      GGCCGTGCTG  AGCGAGGACT  CTCGGGCCCG  AGTGGGTGTG  GCCACATACA  GCAGGGAGCT   1260
        GCTGGTGGCG  GTGCCTGTGG  GGGAGTACCA  GGATGTGCCT  GACCTGGTCT  GGAGCCTCGA   1320
        TGGCATTCCC  TTCCGTGGTG  GCCCCACCCT  GACGGGCAGT  GCCTTGCGGC  AGGCGGCAGA   1380
        GCGTGGCTTC  GGGAGCGCCA  CCAGGACAGG  CCAGGACCGG  CCACGTAGAG  TGGTGGTTTT   1440
        GCTCACTGAG  TCACACTCCG  AGGATGAGGT  TGCGGGCCCA  GCGCGTCACG  CAAGGGCGCG   1500
30      AGAGCTGCTC  CTGCTGGGTG  TAGGCAGTGA  GGGCGTGCGG  GCAGAGCTGG  AGGAGATCAC   1560
        AGGCAGCCCA  AAGCATGTGA  TGGTCTACTC  GGATCCTCAG  GATCTGTTCA  ACCAAATCCC   1620
        TGAGCTGCAG  GGGAAGCTGT  GCAGCCGGCA  GCGGCCAGGG  TGCCGGACAC  AAGCCCTGGA   1680
        CCTCGTCTTC  ATGTTGGACA  CCTCTGCCTC  AGTAGGGCCC  GAGAATTTTG  CTCAGATGCA   1740
        GAGCTTTGTG  AGAAGCTGTG  CCCTCCAGTT  TGAGGTGAAC  CCTGACGTGA  CACAGGTCGG   1800
35      CCTGGTGGTG  TATGGCAGCC  AGGTGCAGAC  TGCCTTCGGG  CTGGACACCA  AACCCACCCG   1860
        GGCTGCGATG  CTGCGGGCCA  TTAGCCAGGC  CCCCTACCTA  GGTGGGGTGG  GCTCAGCCGG   1920
        CACCGCCCTG  CTGCACATCT  ATGACAAAGT  GATGACCGTC  CAGAGGGGTG  CCCGGCCTGG   1980
        TGTCCCCAAA  GCTGTGGTGG  TGCTCACAGG  CGGGAGAGCC  GCAGAGGATG  CAGCCGTTCC   2040
        TGCCCAGAAG  CTGAGGAACA  ATGGCATCTC  TGTCTTGGTC  GTGGGCGTGG  GGCCTGTCCT   2100
40      AAGTGAGGGT  CTGCGGAGGC  TTGCAGGTCC  CCGGGATTCC  CTGATCCACG  TGGCAGCTTA   2160
        CGCCGACCTG  CGGTACCACC  AGGACGTGCT  CATTGAGTGG  CTGTGTGGAG  AAGCCAAGCA   2220
        GCCAGTCAAC  CTCTGCAAAC  CCAGCCCGTG  CATGAATGAG  GGCAGCTGCG  TCCTGCAGAA   2280
        TGGGAGCTAC  CGCTGCAAGT  GTCGGGATGG  CTGGGAGGGC  CCCCACTGCG  AGAACCGATT   2340
        CTTGAGACGC  CCCTGAGGCA  CATGGCTCCC  GTGCAGGAGG  GCAGCAGCCG  TACCCCTCCC   2400
45      AGCAACTACA  GAGAAGGCCT  GGGCACTGAA  ATGGTGCCTA  CCTTCTGGAA  TGTCTGTGCC   2460
        CCAGGTCCTT  AGAATGTCTG  CTTCCCGCCG  TGGCCAGGAC  CACTATTCTC  ACTGAGGGAG   2520
        GAGGATGCC   CAACTGCAGC  CATGCTGCTT  AGAGACAAGA  AAGCAGCTGA  TGTCACCCAC   2580
        AAACGATGTT  GTTGAAAAGT  TTTGATGTGT  AAGTAAATAC  CCACTTTCTG  TACCTGCTGT   2640
        GCCTTGTTGA  GGCTATGTCA  TCTGCCACCT  TTCCCTTGAG  GATAAACAAG  GGGTCCTGAA   2700
        GACTTAAATT  TAGCGGCCTG  ACGTTCCTTT  GCACACAATC  AATGCTGCGC  AGAATGTTGT   2760
50      TGACACAGTA  ATGCCCAGCA  GAGGCCTTTA  CTAGAGCATC  CTTTGGACGG            2760
```

Seq ID NO: 124 Protein sequence
Protein Accession #: Eos sequence

```
        1           11          21          31          41          51
        |           |           |           |           |           |
55      MPPFLLLEAV  CVFLFSRVPP  SLPLQEVEVS  KETIGKISAA  SKMMWCSAAV  DIMFLLDGSN     60
        SVGKGSFERS  KHFAITVCDG  LDISPERVRV  GAFQFSSTFH  LEFPLDSFST  QQEVKAHIKR    120
        MVFKGGHTET  ELALKYLHR   GLPGGRNASV  PQILIIVTDG  KSQGDVALPS  KQLKERGVTV    180
60      PAVGVRFPRW  EKLHALASEP  RGQHVLLAEQ  VEDATNGLFS  TLSSSAICSS  ATPDCRVEAH    240
        PCEHRTLEMV  REFAGNAPCW  RGSRRTLAVL  AAHCPFYSWK  RVFLTHPATC  YRTTCPGPCD    300
        SQPCQNGGTC  VFEGLDGYQC  LCPLAFGGEA  NCALKLSLEC  RVDLLFLLDS  SAGTTLDGFL    360
        RAKVFVKRFV  RAVLSEDSRA  RVGVATYSRE  LLVAVPVGEY  QDVPDLVWSL  DGIPFRGGPT    420
65      LTGSALRQAA  ERGFGSATRT  GQDRPRRVVV  LLTESHSEDE  VAGPARHARA  RELLLIGVGS    480
        EAVRAELEEI  TGSPKHVMVY  SDPQDLFNQI  PELQGKLCSR  QRPGCRTQAL  DLVFMLDTSA    540
        SVGPENFAQM  QSFVRSCALQ  FEVNPDVTQV  GLVVYGSQVQ  TAFGLDTKPT  RAAMLRAISQ    600
        APYLGGVGSA  GTALLHIYDK  VMTVQRGARP  GVPKAVVVLT  GGRGAEDAAV  PAQKLRNNGI    660
        SVLVVGVGPV  LSEGLRRLAG  PRDSLIHVAA  YADLRYHQDV  LIEWLCGEAK  QPVNLCKPSP    720
70      CMNEGSCVLQ  NGSYRCKCRD  GWEGPHCENR  FLRRP
```

Seq ID NO: 125 DNA sequence
Nucleic Acid Accession #: NM_031942.1
Coding sequence: 145..1260

```
        1           11          21          31          41          51
        |           |           |           |           |           |
75      CCCGAGCCCC  GCCCCTCCGG  GCCCGGGTCG  GCGCGCCCAG  CCTGCCAGCC  GCGCTGCTGC     60
        TGCTCCTCCT  GCTGTGGGAC  CGCTGACCGC  GCGGCTGCTC  CGCTCTCCCC  GCTCCAAGCG    120
80      CCGATCTGGG  CACCCGCCAC  CAGCATGGAC  GCTCGCCGCG  TGCCGCAGAA  AGATCTCAGA    180
        GTAAAGAAGA  ACTTAAAGAA  ATTCAGATAT  GTGAAGTTGA  TTTCCATGGA  AACCTCGTCA    240
        TCCTCTGATG  ACAGTTGTGA  CAGCTTTGCT  TCTGATAATT  TTGCAAACAC  GAGGCTGCAG    300
        TCAGTTCGGG  AAGGCTGTAG  GACCCGCAGC  CAGTGCAGGC  ACTCTGGACC  TCTCAGGGTG    360
        GCGATGAAGT  TTCCAGCGCG  GAGTACCAGG  GGAGCAACCA  ACAAAAAAGC  AGAGTCCCGC    420
        CAGCCCTCAG  AGAATTCTGT  GACTGATTCC  AACTCCGATT  CAGAAGATGA  AAGTGGAATG    480
```

```
AATTTTTTGG AGAAAAGGGC TTTAAATATA AAGCAAAACA AAGCAATGCT TGCAAAACTC   540
ATGTCTGAAT TAGAAAGCTT CCCTGGCTCG TTCCGTGGAA GACATCCCCT CCCAGGCTCC   600
GACTCACAAT CAAGGAGACC GCGAAGGCGT ACATTCCCGG GTGTTGCTTC CAGGAGAAAC   660
CCTGAACGGA GAGCTCGTCC TCTTACCAGG TCAAGGTCCC GGATCCTCGG GTCCCTTGAC   720
GCTCTACCCA TGGAGGAGGA GGAGGAAGAG GATAAGTACA TGTTGGTGAG AAAGAGGAAG   780
ACCGTGGATG GCTACATGAA TGAAGATGAC CTGCCCAGAA GCCGTCGCTC CAGATCATCC   840
GTGACCCTTC CGCATATAAT TCGCCCAGTG GAAGAAATTA CAGAGGAGGA GTTGGAGAAC   900
GTCTGCAGCA ATTCTCGAGA GAAGATATAT AACCGTTCAC TGGGCTCTAC TTGTCATCAA   960
TGCCGTCAGA AGACTATTGA TACCAAAACA AACTGCAGAA ACCCAGACTG CTGGGGCGTT  1020
CGAGGCCAGT TCTGTGGCCC CTGCCTTCGA AACCGTTATG GTGAAGAGGT CAGGGATGCT  1080
CTGCTGGATC CGAACTGGCA TTGCCCGCCT TGTCGAGGAA TCTGCAACTG CAGTTTCTGC  1140
CGGCAGCGAG ATGGACGGTG TGCGACTGGG GTCCTTGTGT ATTTAGCCAA ATATCATGGC  1200
TTTGGGAATG TGCATGCCTA CTTGAAAAGC CTGAAACAGG AATTTGAAAT GCAAGCATAA  1260
TATCTGAAAA ATTTGCTGCC TGCCTTCTAC TTCTCAAATC TTTCTTGTAA AAGTTTCCAA  1320
TTTTTTCACT GAAACCTGAG TTAAAAATCT TGATGATCAG CCTGTTTCAT AAGAAACTCC  1380
AATCAAGTTA ATCTTAGCAG ACATGTGTTT CTGGAGCATC ACAGAAGGTA TATTGCTAGT  1440
TACACTTTGC CCTCCTGCAG TTTCTTCTCT GCTCCCAACC CCCATCTCAT AGCATCCCCC  1500
TCTATTTCCA ATGCTCCTCT CCAACCGCTT AGTTTCTGAA TTTCTTTTAA ATTACAGTTT  1560
TATGAAAGCA TATTTTATTT ACTTGGTGTT GAAATAGCCC TCATAAAACC TAAGCACTTG  1620
GAAACACAAT AATAGTATTA ACTAACTAGA TCTATTGAAT TTCAGAGAAG AGCCTTCTAA  1680
CTTGTTTACA CAAAAACGAG TATGATTTAG CACTCATACT AGTTGAAATT TTTAATAGAA  1740
TCAAGGCACA AAAGTCTTAA AACCATGTGG AAAAATTAGG TAATTATTGC AGATTGATGT  1800
CTCTCAATCC CATGTATTGC GCTTATGTTA CAAGTTGTTG TCACAGTTGA GACTTAATTT  1860
CTCCTAATTT CTTCTGCCCG AAGGGTAAGT GGTGCGTCCA GCTTACACGA TCATAATTCA  1920
AAGGTTGGTG GGCAATGTAA TACTTAATTA AAATAATGAT GGAAGAGCTA TCTGGAGATT  1980
ATGAGTAAGC TGATTTGAAT TTTCAGTATA AAACTTTAGT ATAATTGTAG TTTGCAAAGT  2040
TTATTTCAGT TCACATGTAA GGTATTGCAA ATAAATTCTT GGACAATTTT GTATGGAAAC  2100
TTGATATTAA AAACTAGTCT GTGGTTCTTT GCAGTTTCTT GTAAATTTAT AAACCAGGCA  2160
CAAGGTTCAA GTTAGATTT TAAGCACTTT TATAACAATG ATAAGTGCCT TTTTGGAGAT  2220
GTAACTTTTA GCAGTTTGTT AACCTGACAT CTCTGCCAGT CTAGTTTCTG GGCAGGTTTC  2280
CTGTGTCAGT ATTCCCCCTC CTCTTTGCAT TAATCAAGGT ATTTGGTAGA GGTGGAATCT  2340
AAGTGTTTGT ATGTCCAATT TACTTGCATA TGTAAACCAT TGCTGTGCCA TTCAATGTTT  2400
GATGCATAAT TGGACCTTGA ATCGATAAGT GTAAATACAG CTTTTGATCT GTAATGCTTT  2460
TATACAAAAG TTTATTTTAA TAATAAAATG TTTGTTCTAA AAAAAAAAAA         2460
```

Seq ID NO: 126 Protein sequence
Protein Accession #: NP_114148.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MDARRVPQKD LRVKKNLKKF RYVKLISMET SSSSDDSCDS FASDNFANTR LQSVREGCRT   60
RSQCRHSGPL RVAMKFPARS TRGATNKKAE SRQPSENSVT DSNSDSEDES GMNFLEKRAL  120
NIKQNKANLA KLMSELESFP GSFRGRHPLP GSDSQSRRPR RRTFPGVASR RNPERRARPL  180
TRSRSRILGS LDALPMEEER EEDKYMLVRK RKTVDGYMNE DDLPRSRRSR SSVTLPHIIR  240
PVEEITEEEL ENVCSNSREK IYNRSLGSTC HQCRQKTIDT KTNCRNPDCW GVRGQFCGPC  300
LRDNRYGEEVR DALLDPNWHC PPCRGICNCS FCRQRDGRCA TGVLVYLAKY HGFGNVHAYL  360
KSLKQEFEMQ A
```

Seq ID NO: 127 DNA sequence
Nucleic Acid Accession #: AF305616.1
Coding sequence: 1..863

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCACCGCT TGATGGGGGT CAACAGCACC GCGGCGGCCG CCGCCGGGCA GCCCAATGTC   60
TCCTGCACGT GCAACTGCAA ACGCTCTTTG TTCCAGAGCA TGGAGATCAC GGAGCTGGAG  120
TTTGTTCAGA TCATCATCAT CGTGGTGGTG ATGATGGTGA TGGTGGTGGT GATCACGTGC  180
CTGCTGAGCC ACTACAAGCT GTCTGCACGG TCCTTCATCA GCCGGCACAG CCAGGGGCGG  240
AGGAGAGAAG ATGCCCTGTC CTCAGAAGGA TGCCTGTGGC CCTCGGAGAG CACAGTGTCA  300
GGCAACGGAA TCCCAGAGCC GCAGGTCTAC GCCCCGGCCTC GGCCCACCGA CCGCCTGGCC  360
GTGCCGCCCT TCGCCCAGCG GGAGCGCTTC CACCGCTTCC AGCCCACCTA TCCGTACCTG  420
CAGCACGAGA TCGACCTGCC ACCCACCATC TCGCTGTCAG ACGGGGAGGA GCCCCCACCC  480
TACCAGGGCC CCTGCACCCT CCAGCTTCGG GACCCCGAGC AGCAGCTGGA ACTGAACCGG  540
GAGTCGGTGC GCGCACCCCC AAACAGAACC ATCTTCGACA GTGACCTGAT GGATAGTGCC  600
AGGCTGGGCG GCCCCTGCCC CCCCAGCAGT AACTCGGGCA TCAGCGCCAC GTGCTACGGC  660
AGCGGCCGGC GCATGGAGGG GCCGCCGCCC ACCTACAGCG AGGTCATCGG CCACTACCCG  720
GGGTCCTCCT TCCAGCACCA GCAGAGCAGT GGGCCGCCCT CCTTGCTGGA GGGGACCCGG  780
CTCCACCACA CACACATCGC GCCCCTAGAG AGCGCAGCCA TCTGGAGCAA AGAGAAGGAT  840
AAACAGAAAG GACACCCTCT CTAG                                         840
```

Seq ID NO: 128 Protein sequence
Protein Accession #: AAL09357.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MHRLMGVNST AAAAAGQPNV SCTCNCKRSL FQSMEITELE FVQIIIIVVV MMVMVVVITC   60
LLSHYKLSAR SFISRHSQGR RREDALSSEG CLWPSESTVS GNGIPEPQVY APPRPTDRLA  120
VPPFAQRERF HRFQPTYPYL QHEIDLPPTI SLSDGEEPPP YQGPCTLQLR DPEQQLELNR  180
RSVRAPPNRT IFDSDLMDSA RLGGPCPPSS NSGISATCYG SGGRMEGPPP TYSEVIGHYP  240
GSSFQHQQSS GPPSLLEGTR LHHTHIAPLE SAAIWSKEKD KQKGHPL
```

Seq ID NO: 129 DNA sequence
Nucleic Acid Accession #: NM_004952.1

Coding sequence: 1..718

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGCGGCGG CTCCGCTGCT GCTGCTGCTG CTGCTCGTGC CGGTGCCGCT GCTGCCGCTG    60
CTGGCCCAAG GGCCGGAGG GGGCGCTGGGA AACCGGCATG CGGTGTACTG GAACAGCTCC    120
AACCAGCACC TGCGGCGAGA GGGCTACACC GTGCAGGTGA ACGTGAACGA CTATCTGGAT    180
ATTTACTGCC CGCACTACAA CAGCTCGGGG GTGGGCCCCG GGGCGGGACC GGGGCCCGGA    240
GGCGGGGCAG AGCAGTACGT GCTGTACATG GTGAGCCGCA ACGGCTACCG CACCTGCAAC    300
GCCAGCCAGG GCTTCAAGCG CTGGGAGTGC AACCGGCCGC ACGCCCCGCA CAGCCCCATC    360
AAGTTCTCGG AGAAGTTCCA GCGCTACAGC GCCTTCTCTC TGGGCTACGA GTTCCACGCC    420
GGCCACGAGT ACTACTACAT CTCCACGCCC ACTCACAACC TGCACTGGAA GTGTCTGAGG    480
ATGAAGGTGT TCGTCTGCTG CGCCTCCACA TCGCACTCAG GGGAGAAGCC GGTCCCCACT    540
CTCCCCCAGT TCACCATGGG CCCCAATGTG AAGATCAACG TGCTGGAAGA CTTTGAGGGA    600
GAGAACCCTC AGGTGCCCAA GCTTGAGAAG AGCATCAGCG GGACCAGCCC CAAACGGGAA    660
CACCTGCCCC TGGCCGTGGG CATCGCCTTC TTCCTCATGA CGTTCTTGGC CTCCTAG      
```

Seq ID NO: 130 Protein sequence
Protein Accession #: NP_004943.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MAAAPLLLLL LLVPVPLLPL LAQGPGGALG NRHAVYWNSS NQHLRREGYT VQVNVNDYLD    60
IYCPHYNSSG VGPGAGPGPG GGAEQYVLYM VSRNGYRTCN ASQGFKRWEC NRPHAPHSPI    120
KFSEKFQRYS AFSLGYEFHA GHEYYYISTP THNLHWKCLR MKVFVCCAST SHSGEKFVPT    180
LPQFTMGPNV KINVLEDFEG ENPQVPKLEK SISGTSPKRE HLPLAVGIAF FLMTFLAS
```

Seq ID NO: 131 DNA sequence
Nucleic Acid Accession #: NM_012445.1
Coding sequence: 276..1271

```
1          11         21         31         41         51
|          |          |          |          |          |
GCACGAGGGA AGAGGGTGAT CCGACCCGGG GAAGGTCGCT GGGCAGGGCG AGTTGGGAAA    60
GCGGCAGCCC CCGCCGCCCC CGCAGCCCCT TCTCCTCCTT TCTCCCACGT CCTATCTGCC    120
TCTCGCTGGA GGCCAGGCCG TGCAGCATCG AAGACAGGAG GAACTGGAGC CTCATTGGGC    180
GGCCCGGGGC GCCGGCCTCG GGCTTAAATA GGAGCTCCGG GCTCTGGCTG GGACCCGACC    240
GCTGCCGGCC GCGCTCCCGC TGCTCCTGCC GGGTGATGGA AAACCCCAGC CCGGCCGCCG    300
CCCTGGGCAA GGCCCTCTGC GCTCTCCTCC TGGCCACTCT CGGCGCCGCC GGCCAGCCTC    360
TTGGGGGAGA GTCCATCTGT TCCGCCAGAG CCCCGGCCAA ATACAGCATC ACCTTCACGG    420
GCAAGTGGAG CCAGACGGCC TTCCCCAAGC AGTACCCCCT GTTCCGCCCC CCTGCGCAGT    480
GGTCTTCGCT GCTGGGGGCC GCGCATAGCT CCGACTACAG CATGTGGAGG AAGAACCAGT    540
ACGTCAGTAA CGGGCTGCGC GACTTTGCGG AGCGCGGCGA GGGCTGGGCG CTGATGAAGG    600
AGATCGAGCG GGCGGGGGAG GCGCTGCAGA GCGTGCACGC GGGTGTTTTCG GCGGCCCGCGCG  660
TCCCCAGCGG CACCGGGCAG ACGTCGGCGG AGCTGGAGGT GCAGCGCAGG CACTCGCTGG    720
TCTCGTTTGT GGTGCGCATC GTGCCCAGCC CCGACTGGTT CGTGGGCGTG GACAGCCTGG    780
ACCTGTGCGA CGGGGACCGT TGGCGGGAAC AGGCGGCGCT GGACCTGTAC CCCTACGACG    840
CCGGGACGGA CAGCGGCTTC ACCTTCTCCT CCCCCAACTT CGCCACCATC CCGCAGGACA    900
CGGTGACCGA GATAACGTCC TCCTCTCCCA GCCACCGGGC CAACTCCTTC TACTACCCGC    960
GGCTGAAGGC CCTGCCTCCC ATCGCCAGGG TGACACTGGT GCGGCTGCGA CAGAGCCCCA    1020
GGGCCTTCAT CCCTCCCGCC CCAGTCCTGC CCAGCAGGGA CAATGAGATT GTAGACAGCG    1080
CCTCAGTTCC AGAAAACGCCG CTGGACTGCG AGGTCTCCCT GTGGTCGTCC TGGGGACTGT    1140
GCGGGAGGCCA CTGTGGGAGG CTCGGGACCA AGAGCAGGAC TCGCTACGTC CGGGTCCAGC    1200
CCGCCAACAA CGGGAGCCCC TGCCCCGAGC TCGAAGAAGA GGCTGAGTGC GTCCCTGATA    1260
ACTCGGTCTA AGACCAGAGC CCCGCAGCCC CTGGGGCCCC CGGAGCCATG GGGTGTCGGGG   1320
GGCTCCTGTG CAGGCTCATG CTGCAGGCGG CCGAGGCACA GGGGGGTTTCG CGCTGCTCCT   1380
GACCGGCGGTG AGGCCGCGCC GACCATCTCT GCACTGAAGG GCCCTCTGGT GGCCGGCACG   1440
GGCATTGGGA AACAGCCTGG TCCTTTCCCA ACTTTGCTTC TTAGGGGCCC CCGTGTCCCG    1500
TCTGCTCTCA GCCTCCTCCT CCTGCAGGAT AAAGTCATCC CCAAGGCTCC AGCTACTCTA   1560
AATTATGGTC TCCTTATAAG TTATTGCTGC TCCAGGAGAT TGTCCTTCAT CGTCCAGGGG    1620
CCTGGCTCCC ACGTGGTTGC AGATACCTCA GACCTGGTGC TCTAGGCTGT GCTGAGCCCA    1680
CTCTCCCGAG GGCGCATCCA AGCGGGGGCC ACTTGAGAAG TGAATAAATG GGGCGGTTTC    1740
GGAAGCGTCA GTGTTTCCAT GTTATGGATC TCTCTGCGTT TGAATAAAGA CTATCTCTGT    1800
TGCTCAC
```

Seq ID NO: 132 Protein sequence
Protein Accession #:   NP_036577.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MENPSPAAAL GKALCALLLA TLGAAGQPLG GESICBARAP AKYSITFTGK WSQTAFPKQY    60
PLFRPPAQWS SLLGAAHSSD YSMWRKNQYV SNGLRDFAER GEAWALMKEI EAAGEALQSV    120
HAVPSAPAVP SGTGQTSAEL EVQRRHSLVS FVVRIVPSPD WFVGVDSLDL CDGDRWREQA    180
ALDLYPYDAG TDSGFTFSSP NFATIPQDTV TEITSSSPSH PANSFYYPRL KALPPIARVT    240
LVRLRQSPRA FIPPAFVLPS RDNEIVDSAS VPETPLDCEV SLWSSWGLCG GHCGRLGTKS    300
RTRYVRVQPA NNGSPCPELE EEAECVPDNC V
```

Seq ID NO: 133 DNA sequence
Nucleic Acid Accession #: NM_019894
Coding sequence: 1..1314

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGTTACAGG ATCCTGACAG TGATCAACCT CTGAACAGCC TCGATGTCAA ACCCCTGCGC    60
```

```
     AAACCCCGTA TCCCCATGGA GACCTTCAGA AAGGTGGGGA TCCCCATCAT CATAGCACTA   120
     CTGAGCCTGG CGAGTATCAT CATTGTGGTT GTCCTCATCA AGGTGATTCT GGATAAATAC   180
     TACTTCCTCT GCGGGCAGCC TCTCCACTTC ATCCCGAGGA AGCAGCTGTG TGACGGAGAG   240
     CTGGACTGTC CCTTGGGGA GGACGAGGAG CACTGTGTCA AGAGCTTCCC CGAAGGGCCT    300
     GCAGTGGCAG TCCGCCTCTC CAAGGACCGA TCCACACTGC AGGTGCTGGA CTCGGCCACA   360
     GGGAACTGGT TCTCTGCCTG TTTCGACAAC TTCACAGAAG CTCTCGCTGA GACAGCCTGT   420
     AGGCAGATGG GCTACAGCAG CAAACCCAAT TTCAGAGGTG TGGAGATTGG CCCAGACCAG   480
     GATCTGGATG TTGTTGAAAT CACAGAAAAC AGCCAGGAGC TTCGCATGCG GAACTCAAGT   540
     GGGCCTGTC TCTCAGGCTC CCTGGTCTCC CTGCACTGTC TTGCCTGTGG AAGAGCCTG     600
     AAGACCCCC GTGTGGTGGA TGGGGAGGAG GCCTCTGTGG ATTCTTGGCC TTGGCAGGTC    660
     AGCATCCAGT ACGACAAACA GCACGTCTGT GGAGGGAGCA TCCTGGACCC CCACTGGGTC    720
     CTCACGGCAG CCCACTGCTT CAGGGAAACAT ACCGATGTGT TCAACTGGAA GGTGCGGGCA   780
     GGCTCAGACA AACTGGGCAG CTTCCCATCC CTGGCTGTGG CCAAGATCAT CATCATTGAA   840
     TTCAACCCCA TGTACCCCAA AGACAATGAC ATCGCCCTCA TGAAGCTGCA GTTCCCACTC    900
     ACTTTCTCAG GCACAGTCAG GCCCATCTGT CTGCCCTTCT TTGATGAGGA GCTCACTCCA    960
     GCCACCCCAC TCTGGATCAT TGGATGGGGC TTTACGAAGC AGAATGGAGG GAAGATGTCT   1020
     GACATACTGC TGCAGGCGTC AGTCCAGGTC ATTGACAGCA CACGGTGCAA TGCAGACGAT   1080
     GCGTACCAGG GGGAAGTCAC CGAGAAGATG ATGTGTGCAG GCATCCCGGA AGGGGGTGTG   1140
     GACACCTGCC AGGGTGACAG TGGTGGGCCC CTGATGTACC AATCTGACCA GTGGCATGTG   1200
     GTGGGCATCG TTAGCTGGGG CTATGGCTGC GGGGGCCCGA GCACCCCAGG AGTATACACC   1260
     AAGGTCTCAG CCTATCTCAA CTGGATCTAC AATGTCTGGA AGGCTGAGCT GTAA
```

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     MLQDPDSDQP LNSLDVKPLR KPRIPMETFR KVGIPIIIAL LSLASIIIVV VLIKVILDKY   60
     YFLCGQPLHF IPRKQLCDGE LDCPLGEDEE HCVKSFPEGP AVAVRLSKDR STLQVLDSAT  120
     GNWFSACFDN FTEALAETAC RQMGYSSKPT FRAVEIGPDQ DLDVVEITEN SQELRMRNSS  180
     GPCLSGSLVS LHCLACGKSL KTPRVVGGEE ASVDSWPWQV SIQYDKQHVC GGSILDPHWV  240
     LTAAHCFREH TDVFNWKVRA GSDKLGSFPS LAVAKIIIIE FNPMYPKDND IALMKLQPPL  300
     TFSGTVRPIC LPFFDEELTP ATPLWIIGWG FTKQNGGKMS DILLQASVQV IDSTRCNADD  360
     AYQGEVTEKM MCAGIPEGGV DTCQGDSGGP LMYQSDQWHV VGIVSWGYGC GGPSTPGVYT  420
```

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     CGATCCTGCC GGAGCCCGGC CGCCGCCGGC TTGGATTCTG AAACCTTCCT TGTATCCCTC   60
     CTGAGACATC TTTGCTGCAA GATCGAGGCT GTCCTCTGGT GAGAAGGTGG TGAGGCTTCC  120
     CGTCATATTC CAGCTCTGAA CAGCAACATG GGGTGCAAAG TCCTGCTCAA CATTGGGCAG  180
     CAGATGCTGC GGCGGAAGGT GGTGGACTGT AGCCGGGAGG AGACGCGGCT GTCTCGCTGC  240
     CTGAACACTT TTGATCTGGT GGCCCTCGGG GTGGGCAGCA CACTGGGTGC TGGTGTCTAC  300
     GTCCTGGCTG GAGCTGTGGC CCGTGAGAAT GCAGGCCCTG CCATTGTCAT CTCCTTCCTG  360
     ATCGCTGCGC TGGCCTCAGT GCTGGCTGGC CTGTGCTATG GCGAGTTTGG TGCTCGGGTC  420
     CCCAAGACGG GCTCAGCTTA CCTCTACAGC TATGTCACCG TTGGAGAGCT CTGGGCCTTC  480
     ATCACCGGCT GGAACTTAAT CCTCTCCTAC ATCATCGGTA CTTCAAGCGT AGCGAGGGCC  540
     TGGAGCGCCA CCTTCGACGA GCTGATAGGC AGACCCATCG GGGAGTTCTC GGGCCCTCTG  600
     ATGACTCTGA ACGCCCCCGG CGTGCTGGCT GAAAACCCCG ACATATTCGC AGTGATCATA  660
     ATTCTCATCT TGACAGGACT TTTAACTCTT GGTGTGAAAG AGTCGGCCAT GGTCAACAAA  720
     ATATTCACTT GTATTAACGT CCTGGTCCTG GGCTTCATAA TGGTGTCAGG ATTTGTGAAA  780
     GGATCGGTTA AAAACTGGCA GCTCACGGAG GAGGATTTTG GGAACACATC AGGCCGGTCTC  840
     TGTTTGAACA ATGACACAAA AGAAGGGAAG CCCGGTGTTG GTGGATTCAT GCCCTTCGGG  900
     TTCTCTGGTG TCCGTGTCGGG GGCAGCGACT TGCTTCATAA CCTTCGTCGG CTTTGACTGC  960
     ATCGCCACCA CAGGTGAAGA GGTGAAGAAC CCACAGAAGG CCATCCCCGT GGGGATCGTG  1020
     GCGTCCCTCT TGATCTGCTT CATCGCCTAC TTTGGGGTGT CGGCTGCCCT CACGGCTCATG  1080
     ATGCCCTACT TCTGCCTGGA CAATAACAGC CCCCTGCCCG ACGCCTTTAA GCACGTGGGC  1140
     TGGGAAGGTG CCAAGTACGC AGTGGCCGTG GGCTCCCTCT GCGCTCTTTC CGCCAGTCTT  1200
     CTAGGTTCCA TGTTTCCCAT GCCTCGGGTT ATCTATGCCA TGGCTGAGGA TGGACTGCTA  1260
     TTTAAATTCT TAGCCAACGT CAATGATAGG ACCAAAACAC CAATAATCGC CAATTAGCC   1320
     TCGGGTGCCG TTGCTGCTGT GATGGCCTTC CTCTTTGACC TGAAGGACTT GGTGGACCTC  1380
     ATGTCCATTG GCACTCTCCT GGCTTACTCG TTGGTGGCTG CCTGTGTGTT GGTCTTACGG  1440
     TACCAGCCAG AGCAGCCTAA CCTGGTATAC CAGATGGCCA GTACTTCCGA CGAGTTAGAT  1500
     CCAGCAGACC AAAATGAATT GGCAAGCACC AATGATTCCC AGCTGGGGTT TTTACCAGAG  1560
     GCAGAGATGT TCTCTTTGAA AACCATACTC TCACCCAAAA ACATGGAGCC TTCCAAAATC  1620
     TCTGGGCTAA TTGTGAACAT TTCAACCAGC CTTATAGCTG TTCTCATCAT CACCTTCTGC  1680
     ATTGTGACCG TGCTTGGAAG GGAGGCTCTC ACCAAAGGGG CGCTGTGGGC AGTCTTTCTG  1740
     CTCGCAGGGT CTGCCCTCCT CTGTGCCGTG GTCACGGGCG TCATCTGGAG GCAGCCCGAG  1800
     AGCAAGACCA AGCTCTCATT TAAGGTTCCC TTCCTGCCAG TGCTCCCCAT CCTGAGCATC  1860
     TTCGTGAACG TCTATCTCAT GATGCAGCTG GACCAGGGCA GTTTGCTGTG  1920
     TGGATGCTGA TAGGCTTCAT CATCTACTTT GGCTATGGCC TGTGGCACAG CGAGGAGGCG  1980
     TCCCTGGATG CCGACCAAGC AAGGACTCCT GACGGCAACT TGGACCAGTG CAAGTGACGC  2040
     ACAGCCCGGC CCCCCGGAGG TGGCAGCAGC CCCGAGGGAC GCCCCAGAG GACGGGGAGG   2100
     CACCCCCACCC TCCCCACCAG TGCAACAGAA ACCACCTGCC TCCACACCCT CACTGCA
```

```
     1          11         21         31         41         51
     |          |          |          |          |          |
```

```
      MGCKVLLNIG QQMLRRKVVD CSREETRLSR CLNTFDLVAL GVGSTLGAGV YVLAGAVARE    60
      NAGPAIVISF LIAALASVLA GLCYGBFGAR VPKTGSAYLY SYVTVGELWA FITGWNLILS   120
      YIIGTSSVAR AWSATFDELI GRPIGEFSRT HMTLNAPGVL AENPDIPAVI IILILTGLLT   180
      LGVKESAMVN KIFTCINVLV LGFIMVSGFV KGSVKNWQLT EEDFGNTSGR LCLNNDTKEG   240
      KPGVGGFMPF GFSGVLSGAA TCFYAFVGFD CIATTGEEVK NPQKAIPVGI VASLLICFIA   300
      YFGVSAALTL MMPYFCLDNN SPLPDAFKHV GWEGAKYAVA VGSLCALSAS LLGSMPPMPR   360
      VIYAMAEDGL LFKFLANVND RTKTPIIATL ASGAVAAVMA FLFDLKDLVD LMSIGTLLAY   420
      SLVAACVLVL RYQPEQPNLV YQMASTSDEL DPADQNELAS TNDSQLGFLP EAEMFSLKTI   480
      LSPKNMEPSK ISGLIVNIST SLIAVLIITF CIVTVLGREA LTKGALWAVF LLAGSALLCA   540
      VVTGVIWRQP ESKTKLSFKV PPLPVLPILS IFVNVYLMMQ LDQGTWVRFA VWMLIGFIIY   600
      FGYGLWHSEE ASLDADQART PDGNLDQCR
```

Seq ID NO: 137 DNA sequence
Nucleic Acid Accession #: NM_032044.1
Coding sequence: 182..658

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      AAGATATAAA AGCTCCAGAA ACGTTGACTG GGACCACTGG AGACACTGAA GAAGGCAGGG    60
      GCCCTTAGAG TCTTGGTTGC CAAACAGATT TGCAGATCAA GGAGAACCCA GGAGTTTCAA   120
      AGAAGCGCTA GTAAGGTCTC TGAGATCCTT GCACTAGCTA CATCCTCAGG GTAGGAGGAA   180
      GATGGCTTCC AGAAGCATGC GGCTGCTCCT ATTGCTGAGC TGCCTGGCCA AAACAGGAGT   240
      CCTGGGTGAT ATCATCATGA GACCCAGCTG TGCTCCTGGA TGGTTTACC ACAAGTCCAA   300
      TTGCTATGGT TACTTCAGGA AGCTGAGGAA CTGGTCTGAT GCCGAGCTCG AGTGTCAGTC   360
      TTACGGAAAC GGAGCCCACC TGGCATCTAT CCTGAGTTTA AAGGAAGCCA GCACCCATAGC   420
      AGAGTACATA AGTGGCTATC AGAGAAGCCA GCCGATATGG ATTGGCCTGC ACGACCCACA   480
      GAAGAGGCAG CAGTGGCAGT GGATTGATGG GGCCATGTAT CTGTACAGAT CCTGGTCTGG   540
      CAAGTCCATG GGTGGGAACA AGCACTGTGC TGAGATGAGC TCCAATAACA ACTTTTTAAC   600
      TTGGAGCAGC AACGAATGCA ACAAGCGCCA ACACTTCCTG TGCAAGTACC GACCATAGAG   660
      CAAGAATCAA GATTCTGCTA ACTCCTGCAC AGCCCGTCC TCTTCCTTTC TGCTAGCCTG   720
      GCTAAATCTG CTCATTATTT CAGAGGGGAA ACCTAGCAAA CTAAGAGTCA TAAGGGCCCT   780
      ACTACACTGG CTTTTTAGG CTTAGAGACA GAAACTTTAG CATTGGCCCA GTAGTGGCTT   840
      CTAGCTCTAA ATGTTTGCCC CGCCATCCCT TTCCACAGTA TCCTTCTTCC CTCCTCCCCT   900
      GTCTCTGGCT GTCTCGAGCA GTCTAGAAGA GTGCATCTCC AGCCTATGAA ACAGCTGGGT   960
      CTTTGGCCAT AAGAAGTAAA GATTTGAAGA CAGAAGGAAG AAACTCAGGA GTAAGCTTCT  1020
      AGACCCCTTC AGCTTCTACA CCCTTCTGCC CTCTCTCCAT TGCCTGCACC CCACCCCAGC  1080
      CACTCAACTC CTGCTTGTTT TTCCTTTGGC CATAGGAAGG TTTACCAGTA GAATCCTTGC  1140
      TAGGTTGATG TGGGCCATAC ATTCCTTTAA TAAACCATTG TGTACATAAG AAAAAAAAAA  1140
```

Seq ID NO: 138 Protein sequence
Protein Accession #:   NP_114433.1

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      MASRSMRLLL LLSCLAKTGV LGDIIMRPSC APGWFYHKSN CYGYFRKLRN WSDAELECQS    60
      YGNGAHLASI LSLKEASTIA EYISGYQRSQ PIWIGLHDPQ KRQQWQWIDG AMYLYRSWSG   120
      KSMGGNKHCA EMSSNNNFLT WSSNECNKRQ HFLCKYRP
```

Seq ID NO: 139 DNA sequence
Nucleic Acid Accession #: XM_051860.2
Coding sequence: 52..3042

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      GCTCACCCAG GAAAAATATG CAATCGTCCC ATTGATATAC AGGCCACTAC AATGGATGGA    60
      GTTAACCTCA GCACCGAGGT TGTCTACAAA AAAGGCCAGG ATTATAGGTT TGCTTGCTAC   120
      GACCGGGGCA GAGCCTGCCG GAGCTACCGT GTACGGTTCC TCTGTGGGAA GCCTGTGAGG   180
      CCCAAACTCA CAGTCACCAT TGACACCAAT GTGAACAGCA CCATTCTGAA CTTGGAGGAT   240
      AATGTACAGT CATGGAAACC TGGAGATACC CTGGTCATTG CCAGTACTGA TTACTCCATG   300
      TACCAGGCAG AAGAGTTCCA GGTGCTTCCC TGCAGATCCT GCGCCCCCAA CCAGGTCAAA   360
      GTGGCAGGGA AACCAATGTA CCTGCACATC GGGGAGGAGA TAGACGGCGT GGACATGCGG   420
      GCGGAGGTTG GGCTTCTGAG CCGGAACATC ATAGTGATGG GGGAGATGGA GGACAAATGC   480
      TACCCCTACA GAAACCACAT CTGCAATTTC TTTGACTTCG ATACCTTTGG GGGCCACATC   540
      AAGTTTGCTC TGGGATTTAA GGCAGCACAC TTGGAGGGCA CGGAGCTGAA GACTATGGGA   600
      CAGCAGCTGG TGGGTCAGTA CCCGATTCAC TTCCACCTGG CCGGTGATGT AGACGAAAGG   660
      GGAGGTTATG ACCCCACCCAC ATACATCAGG GACCTCTCCA TCCATCATAC ATTCTCTCGC   720
      TGCGTCACAG TCCATGGCTC CAATGGCTTG TTGATCAAGG ACGTTGTTGGG CTATAACTCT   780
      TTGGGCCACT GCTTCTTCAC GGAAGATGGG CCGGAGGAAC GCAACACTTT TGACCACTGT   840
      CTTGGCCTCC TTGTCAAGTC TGGAACCCTC CTCCCCTCGG ACCGTGACAG CAAGATGTGC   900
      AAGATGATCA CAGGAGACTC CTACCCAGGG TACATCCCCA AGCCCAGGCA AGACTGCAAT   960
      GCTGTGTCCA CCTTCTGGAT GGCCAATCCC AACAACAACC TCATCAACTG TGCCGCTGCA  1020
      GGATCTGAGG AAACTGGATT TTGGTTTATT TTTCACCACG TACCAACGGG CCCCTCCGTG  1080
      GGAATGTACT CCCCAGGTTA TTCAGAGCAC ATTCCACTGG GAAAATTCTA TAACAACCGA  1140
      GCACATTCCA ACTACCGGCA TGGCATGATC ATAGACAACG GAGTCAAAAC CGAGGAGGCC  1200
      TCTGCCAAGG ACAAGCGGCC GTTCCTCTCA ATCATCTCTG CCAGATACAG CCCTCACCAG  1260
      GACGCCGACC CGCTCGAAGCC CCGGGGAGCCG GCCATCATCA GACACTTCAT TGCCTACAAG  1320
      AACCAGGACC ACGGGGCCTG GCTGCGCCGGC GGGGATGTGT GGCTGGACAG CTGCCGGTTT  1380
      GCTGACAATG GCATTGGCCT GACCCCTGGCC AGTGGTGGAA CCTTCCGGTA TGACTACGGC  1440
      TCCAAGCAAG AGATAAAGAA CAGCTTGTTT GTTGGCGAGA GTGGCAACGT GGGGACGGAA  1500
      ATGATGACA ATAGGATCTG GGGCCCTGGC GGCTTGGACC ATAGCGGAAG GACCCTCCCT  1560
      ATAGGCCAGA ATTTTCCAAT TAGAGGAATT CAGTTATATG ATGGCCCCAT CAACATCCCA  1620
      AACTGCACTT TCCGAAAGTT TGTGGCCCTG GAGGGCCGGC ACACCAGCGC CCTGGCCTTC  1680
      CGGCTGAATA ATGCCTGGCA GAGCTGCCCC CATAACAACG TGACGGGCAT TGCCTTTGAG  1740
      GACGTTCCGA TTACTTCCAG AGTGTTCTTC GGAGAGCCTG GGCCCTGGTT CAACCAGCTG  1800
```

```
GACATGGATG GGGATAAGAC ATCTGTGTTC CATGACGTCG ACGGCTCCGT GTCCGAGTAC 1860
CCTGGCTCCT ACCTCACGAA GAATGACAAC TGGCTGGTCC GGCACCCAGA CTGCATCAAT 1920
GTTCCCGACT GGAGAGGGGC CATTTGCAGT GGGTGCTATG CACAGATGTA CATTCAAGCC 1980
TACAAGACCA GTAACCTGCG AATGAAGATC ATCAAGAATG ACTTCCCCAG CCACCCTCTT 2040
TACCTGGAGG GGGCGCTCAC CAGGAGCACC CATTACCAGC AATACCAACC GGTTGTCACC 2100
CTGCAGAAGG GCTACACCAT CCACTGGGAC CAGACGGCCC CCGCCGAACT CGCCATCTGG 2160
CTCATCAACT TCAACAAGGG CGACTGGATC CGAGTGGGGC TCTGCTACCC GCGAGGCACC 2220
ACATTCTCCA TCCTCTCGGA TGTTCACAAT CGCCTGCTGA AGCAAACGTC CAAGACCGGC 2280
GTCTTCGTGA GGACCTTGCA GATGGACAAA GTGGAGCAGA GCTACCCTGG CAGGAGCCAC 2340
TACTACTGGG ACGAGGACTC AGGGCTGTTG TTCCTGAAGC TGAAAGCTCA GAACGAGAGA 2400
GAGAAGTTTG CTTTCTGCTC CATGAAAGGC TGTGAGAGGA TAAAGATTAA AGCTCTGATT 2460
CCAAAGAACG CAGGCGTCAG TGACTGCACA GCCACAGCTT ACCCCAAGTT CACCGAGAGG 2520
GCTGTCGTAG ACGTGCCGAT GCCCAAGAAG CTCTTTGGTT CTCAGCTGAA AACAAAGGAC 2580
CATTTCTTGG AGGTGAAGAT GGAGAGTTCC AAGCAGCACT TCTTCCACCT CTGGAACGAC 2640
TTCGCTTACA TTGAAGTGGA TGGGAAGGAAG TACCCCAGTT CGGAGGATGG CATCCAGGTG 2700
GTGGTGATTG ACGGGAACCA AGGGCGCGTG GTGAGCCACA CGAGCTTCAG GAACTCCATT 2760
CTGCAAGGCA TACCATGGCA GCTTTTCAAC TATGTGGCGA CCATCCCTGA CAATTCCATA 2820
GTGCTTATGG CATCAAAGGG AAGATACGTC TCCAGAGGCC CATGGACCAG AGTGCTGGAA 2880
AAGCTTGGGG CAGACAGGGG TCTCAAGTTG AAAGAGCAAA TGGCATTCGT TGGCTTCAAA 2940
GGCAGCTTCC GGCCCATCTG GGTGACACTG GACACTGAGG ATCACAAAGC CAAAATCTTC 3000
CAAGTTGTGC CCATCCCTGT GGTGAAGAAG AAGAAGTTGT GAGGCACAGCT GCCGCCCCGT 3060
GCCACCTCGT GGTAGACTAT GACGGTGACT CTTGGCAGCA GACCAGTGGG GGATGGCTGG 3120
GTCCCCAGC CCCTGCCAGC AGCTGCCTGG GAAGGCCGTG TTTCAGCCCT GATGGGCCAA 3180
GGGAAGGCTA TCAGAGACCC TGGTGCTGCC ACCTGCCCCT ACTCAAGTGT CTACCTGGAG 3240
CCCCTGGGGC GGTGCTGGCC AATGCTGGAA ACATTCACTT TCCTGCAGCC TCTTGGGTGC 3300
TTCTCTCCTA TCTGTGCCTC TTCAGTGGGG GTTTGGGGAC CATATCAGGA GACCTGGGTT 3360
GTGCTGACAG CAAAGATCCA CTTTGGCAGG AGCCCTGACG AGCCCTAGGAG GTAGTCTGGA 3420
GGGCTGGTCA TTCACAGATC CCCATGGTCT TCAGCAGACA AGTGAGGGTG GTAAATGTAG 3480
GAGAAAGAGC CTTGGCCTTA AGGAAATCTT TACTCCTGTA AGCAAGAGCC AACCTCACAG 3540
GATTAGGAGC TGGGGTAGAA CTGGCTATCC TTGGGGAAGA GGCAAGCCCT GCCTCTGGCC 3600
GTGTCCACCT TTCAGGAGAC TTTGAGTGGC AGGTTTGGAC TTGGACTAGA TGACTCTCAA 3660
AGGCCCTTTT AGTTCTGAGA TTCCAGAAAT CTGCTGCATT TCACATGGTA CCTGGAACCC 3720
AACAGTTCAT GGATATCCAC TGATATCCAT GATGCTGGGT GCCCCAGCGC ACACGGGATG 3780
GAGAGGTGAG AACTAATGCC TAGCTTGAGG GGTCTGCAGT CCAGTAGGGC AGGCAGTCGA 3840
GTCCATGTGC ACTGCAATGC CAGGTGGAGA AATCACAGAG AGGTAAAATG GAGGCCAGTG 3900
CCATTTCACA GGGGAGGCTC AGGAAGGCTT CTTGCTTACA GGAATGAAGG CTGGGGGCAT 3960
TTTGCTGGGG GGAGATGAGG CAGCCTCTGG AATGGCTCAG GGATTCAGCC CTCCCTGCCG 4020
CTGCCTGCTG AAGCTGGTGA CTACGGGGTC GCCCTTTGCT CACGTCTCTC TGGCCCACTC 4080
ATGATGGAGA AGTGTGGTCA GAGGGGGAGCA ATGGGCTTTG CTGCTTATGA GCACAGAGGA 4140
ATTCAGTCCC CAGGCAGCCC TGCCTCTGAC TCCAAGAGGG TGAAGTCCAC AGAAGTGAGC 4200
TCCTGCCTTA GGGCCTCATT TGCTCTTCAT CCAGGGAACT GAGCACAGGG GGCCTCCAGG 4260
AGACCCTAGA TGTGCTCGTA CTCCCTCGGC CTGGGATTTC AGAGCTGGAA ATATAGAAAA 4320
TATCTAGCAC AAAGCCTTCA TTTTAACAGA TGGGGAAAGT GAGCCCCCAA GATGGGAAAG 4380
AACCACACAG CTAAGGGAGG GCCTGGGGAG CCCCACCCTA GCCCTTGCTG CCACACCACA 4440
TTGCCTCAAC AACCGGCCCC AGAGTGCCCA GGCACTCCTG AGGTAGCTTC TGGAAATGGG 4500
GACAAGTCCC CTCGAAGGAA AGGAAATGAC TAGAGTAGAA TGACAGCTAG CAGATCTCTT 4560
CCCTCCTGCT CCCAGCGCAC ACAAACCCGC CCTCCCCTTG GTGTTGGCGG TCCCTGTCGC 4620
CTTCACTTTG TTCACTACCT GTCAGCCCAG CCTGGGTGCA CAGTAGCTGC AACTCCCCAT 4680
TGGTGCTACC TGGCTCTCCT GTCTCTGCAG CTCTACAGGT GAGGCCCAGC AGAGGGAGTA 4740
GGGGCTCGCCA TGTTTCTGGT GAGCCAATTT GGCTGATCTT GGGTGTCTGA ACAGCTATTG 4800
GGTCCACCCC AGTCCCTTTC AGCTGCTGCT TAATGCCCTG CTCTCTCCCT GGCCCACCTT 4860
ATAGAGACCA CAAAGAGCTC CTGTAAGAGG GAGAACTCTA TCTGTGGTTT ATAATCTTGC 4920
ACGAGGCACC AGAGTCTCCC TGGGTCTTGT GATGAACTAC ATTTATCCCC TTTCCTGCCC 4980
CAACCACAAA CTCTTTCCTT CAAAGAGGGC CTGCCTGGCT CCCTCCACCC AACTGCACCC 5040
ATGAGAACTCG GTCCAAGAGT CCATTCCCCA GGTGGGAGCC AACTGTCAGG GAGGTCTTTC 5100
CCACCAAACA TCTTTCAGCT GCTGGGAGGT GACCATAGGG CTCTGCTTTT AAAGATATGG 5160
CTGCTTCAAA GGCCAGAGTC ACAGGAAGGA CTTCTTCCAG GGAGATTAGT GGTGATGGAG 5220
AGGAGAGTTA AAATGACCTC ATGTCCTTCT TGTCCACGGT TTTGTTGAGT TTTCACTCTT 5280
CTAATGCAAG GGTCTCACAC TGTGAACCAC TTAGGATGTG ATCACTTTCA GGTGGCCAGG 5340
AATGTTGAAT GTCTTTGGCT CAGTTCATTT AAAAAAGATA TCTATTTGAA AGTTCTCAGA 5400
GTTGTACATA TGTTTCACAG TACAGGATCT GTACATAAAA GTTTCTTTCC TAAACCATTC 5460
ACCAAGAGCC AATATCTAGG CATTTTCTTG GTAGCACAAA TTTTCTTATT GCTTAGAAAA 5520
TTGTCCTCCT TGTTATTTCT GTTTGTAAGA CTTAAGTGAG TTAGGTCTTT AAGGAAAGCA 5580
ACGCTCCTCT GAAATGCTTG TCTTTTTTCT GTTGCCGAAA TAGCTGGTCC TTTTTCGGGA 5640
GTTAGATGTA TAGAGTGTTT GTATGTAAAC ATTTCTTGTA GGCATCACCA TGAACAAAGA 5700
TATATTTTCT ATTTATTTAT TATATGTGCA CTTCAAGAAG TCACTGTCAG AGAAATAAAG 5760
AATTGTCTTA AATGTCAAAA AAAAAAAAAA AAAAAAAAA AAAAAAAA
```

Seq ID NO: 140 Protein sequence
Protein Accession #: XP_051860.2

```
1           11          21          31          41          51
|           |           |           |           |           |
MDGVNLSTEV  VYKKGQDYRF  ACYDRGRACR  SYRVRFLCGK  PVRPKLTVTI  DTNVNSTILN    60
LEDNVQSWKP  GDTLVIASTD  YSMYQAEEFQ  VLPCRSCAPN  QVKVAGKPMY  LHIGEEIDGV   120
DMRABVGLLS  RNIIVMGEME  DKCYPYRNHI  CNPFDFDTPG  GHIKFALGFK  AAHLEGTELK   180
HMGQQLVGQY  PIHFHLAGDV  DERGGYDPPT  YIRDLSIEHT  PSRCVTVHGS  NGLLIKDVVG   240
YNSLGHCFFT  PDGPEERNTF  DHCLGLLVKS  GTLLPSDRDS  KMCKMITGDS  YPGYIPKPRQ   300
DCNAVSTFWM  ANPNNNLINC  AAAGSEETGF  WFIFHHVPTG  PSVGMYSPGY  SEHIPLGKFY   360
NNRAHSNYRA  GMIIDNGVKT  TEASAKDKRP  FLSIISARYS  PHQDADPLKP  REPAIIRHFI   420
AYKNQDHGAW  LRGGDVWLDS  CRFADNGIGL  TLASGGTFPY  DDGSKQEIKN  SLFVGESGNV   480
GTEMMDNRIW  GPGGLDHSGR  TLPIGQNFPI  RGIQLYDGPI  NIQNCTFRKF  VALEGRHTSA   540
LAFRLNNAWQ  SCPHNNVTGI  AFEDVPITSR  VFPGEPGPWF  NQLDMDGDKT  SVPHDVDGSV   600
```

```
     SEYPGSYLTK NDNWLVRHPD CINVPDWRGA ICSGCYAQMY IQAYKTSNLR MKIIKNDFPS  660
     HPLYLEGALT RSTHYQQYQP VVTLQKGYTI HWDQTAPAEL AIWLINPNKG DWIRVGLCYP  720
     RGTFFSILSD VHNRLLKQTS KTGVFVRTLQ MDKVEQSYPG RSHYYWDEDS GLLFLKLKAQ  780
     NEREKFAFCS MKGCERIKIK ALIPKNAGVS DCTATAYPKF TERAVVDVPH PKKLPGSQLK  840
     TKDHFLEVKM ESSKQHFFHL WNDPAYIEVD GKKYPSSEDG IQVVVIDGNQ GRVVSETSFR  900
     NSILQGIPWQ LFNYVATIPD NSIVLMASKG RYVSRGPWTR VLEKLGADRG LKLKEQMAFV  960
     GPKGSFRPIW VTLDTEDHKA KIFQVVPIPV VKKKKL
```

Seq ID NO: 141 <u>DNA sequence</u>
Nucleic Acid Accession #: Eos sequence
Coding sequence: 261..2861

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     GAGCTAGCGC TCAAGCAGAG CCCAGCGCGG TGCTATCGGA CAGAGCCTGG CGAGCGCAAG   60
     CGGCBCGGGG AGCCAGCGGG GCTGAGCGCG GCCAGGGTCT GAACCCAGAT TTCCCGAGACT  120
     AGCTACCACT CCGCTTGCCC ACGCCCGGG AGCTCGCGGC GCCTGGCGGT CAGCGACCAG    180
     ACGTCCGGGG CCGCTGCGCT CCTGGCCCGC GAGGCGTGAC ACTGTCTCGG CTACAGACCC   240
     AGAGGGAGCA CACTGCCAGG ATGGGAGCTG CTGGGAGGCA GGACTTCCTC TTCAAGGCCA   300
     TGCTGACCAT CAGCTGGCTC ACTCTGACCT GCTTCCCTGG GGCCACATCC ACAGTGGCTG   360
     CTGGGTGCCC TGACCAGAGC CCTGAGTTGC AACCCTGGAA CCCTGGCCAT GACCAAGACC   420
     ACCATGTGCA TATCGGCCAG GGCAAGACAC TGCTGCTCAC CTCTTCTGCC ACGGTCTATT   480
     CCATCCACAT CTCAGAGGGA GGCAAGCTGG TCATTAAAGA CCACGACGAG CCGATTGTTT   540
     TGCGAACCCG GCACATCCTG ATTGACAACG GAGGAGAGCT GCATGCTGGG AGTGCCCTCT   600
     GCCCTTTCCA GGGCAATTTC ACCATCATTT TGTATGGAAG GGCTGATGAA GGTATTCAGC   660
     CGGATCCTTA CTATGGTCTG AAGTACATTG GGGTTGGTAA AGGAGGCGCT CTTGAGTTGC   720
     ATGGACAGAA AAAGCTCTCC TGGACATTTC TGAACAAGAC CCTTCACCCA GGTGGCATGG   780
     CAGAAGGAGG CTATTTTTTT GAAAGGAGCT GGGGCCACCG TGGAGTTATT GTTCATGTCA   840
     TCGACCCCAA ATCAGGCACA GTCATCCATT CTGACCGGTT TGACACCTAT AGATCCAAGA   900
     AAGAGAGTGA ACGTCTGGTC CAGTATTTGA ACGCGGTGCC CGATGGCAGG ATCCTTTCTG   960
     TTGCAGTGAA TGATGAAGGT TCTCGAAATC TGGATGACAT GGCCAGGAAG GCGATGACCA  1020
     AATTGGGAAG CAAACACTTC CTGCACCTTG GATTTAGACA CCCTTGGAGT TTTCTAACTG  1080
     TGAAAGGAAA TCCATCATCT TCAGTGGAAG ACCATATTGA ATATCATGGA CATCGAGGCT  1140
     CTGCTGCTGC CCGGGTATTC AAATTGTTCC AGACAGAGCA TGGCGAATAT TTCAATGTTT  1200
     CTTTGTCCAG TGAGTGGGTT CAAGACGTGG AGTGGACGGA GTGGTTCGAT CATGATAAAG  1260
     TATCTCAGAC TAAAGGTGGG GAGAAAATTT CAGACCTCTG GAAAGCYCAC CCAGGAAAAA  1320
     TATGCAATCG TCCCATTGAT ATACAGGCCA CTACAATGGA TGGAGTTAAC CTCAGCACCG  1380
     AGGTTGTCTA CAAAAAGGC CAGGATTATA GGTTGCTTG CTACGACCGG GGCAGAGCCT  1440
     GCCGGGAGCTA CCGTGTACGG TTCCTCTGTG GGAAGCCTGT GAGGCCCAAA CTCACAGTCA  1500
     CCATTGACAC CAATGTGAAC AGCACCATTC TGAACTTGGA GGATAATGTA CAGTCATGGA  1560
     AACCTGGAGA TACCCTGGTC ATTGCCAGTA CTGATTACTC CATGTACCAG GCAGAAGAGT  1620
     TCCAGGTGCT TCCCTGCAGA TCCTGCGCCC CCAACCAGGT CAAAGTGGCA GGGAAACCAA  1680
     TGTACCTGCA CATCGGGGAG GAGATAGACG GCGTGGACAT GCGGGGCGAG GTTGGGCTTC  1740
     TGAGCCGGAA CATCATAGTG ATGGGGGAGA TGGAGGACAA ATGCTACCCC TACAGAAACC  1800
     ACATCTGCAA TTTCTTTGAC TTCGATACCT TTGGGGGCCA CATCAAGTTT GCTCTGAACG  1860
     TTAAGGCAGC ACACTTGGAG GGCACGGGAGC TGAAGCATAT GGGACAGCAG CTGGTGGGTC  1920
     AGTACCGAT TCACTTCCAC CTGGCCGGTG ATGTAGACGA AAGGGGAGGT TATGACCCAC  1980
     CCACATACAT CAGGGACCTC TCCATCCATC ATACATTCTC TCGCTGCGTC ACAGTCCATG  2040
     GCTCCAATGG CTTGTTGATC AAGGACGTTG TGGGCTATAA CTCTTTGGGC CACTGCTTCT  2100
     TCACGGAAGA TGGGCCGGAG GAACGCAACA CTTTTGACCA CTGTCTTGGC CTCCTTGTCA  2160
     AGTCTGGAAC CCTCCTCCCC TCGGACCGTG ACAGCAAGAT GTGCAAGATG ATCACGAGAGG  2220
     ACTCCTACCC AGGGTACATC CCCAAGCCCA GGCAAGACTG CAATGCTGTG TCCACCTTCT  2280
     GGATGGCCAA TCCCAACAAC AACCTCATCA ACTGTGCCGC TGCAGGATCT GAGGGAAACTG  2340
     GATTTTTGGTT TATTTTTCAC CACGTACCAA CGGGCCCCTC CGTGGGAATG TACTCCCCAG  2400
     GTTATTCAGA GCACATTCCA CTGGGAAAAT TCTATAACAA CCGAGCACAT TCCAACTACC  2460
     GGGCTGGCAT GATCATAGAC AACGGAGTCA AAACCACCGA GGCCTCTGCC AAGGACAAGC  2520
     GGCCGTTCCT CTCAATCATC TCTGCCAGAT ACAGCCCTCA CCAGGACGCC GACCCGCTGA  2580
     AGCCCCGGGA GCCGGCCATC ATCAGACACT TCATTGCCTA CAAGAACCAG GACCACGGGG  2640
     CCTGGCTGCG CGGCGGGGAT GTGTGGCTGG ACAGCTGCCA TTTCAGAGGG GAGGCTCAGG  2700
     AAGGCTTCTT GCTTACAGGA ATGAAGGCTG GGGGCATTTT GCTGGGGGGA GATGAGGCAG  2760
     CCTCTGGAAT GGCTCAGGGA TTCAGCCCTC CCTGCCGCTG CCTGCTGAAG CTGGTGACTA  2820
     CGGGGTCGCC CTTTGCTCAC GTCTCTCTGG CCCACTCATG ATGGAGAAGT GTGGTCAGAG  2880
     GGGAGCAATG GGCTTTGCTG CTTATGGAGCA CAGAGGAATT CAGTCCCCAG GCAGCCCTGC  2940
     CTCTGACTCC AAGAGGGTGA AGTCCACAGA AGTGAGCTCC TGCCTTAGGG CCTTCATTTGC  3000
     TCTTCATCCA GGGAACTGAG CACAGGGGGC CTCCAGGAGA CCCTAGATGT GCTCGTACTC  3060
     CCTCGGCCTG GGATTTCAGA GCTGGAAATA TAGAAAATAT CTAGCCCAAA GCCTTCATTT  3120
     TAACAGATGG GGAAAGTGAG CCCCCAAGAT GGGAAAGAAC CACACAGCTA AGGGAGGGCC  3180
     TGGGGAGCCC CACCCTAGCC CTTGCTGCCA CACCACATTG CCTCAACAAC CGGCCCCAGA  3240
     GTGCCCAGGC ACTCCTGAGG TAGCTTCTGG AAATGGGGAC AAGTCCCTC GAAGGAAAGG  3300
     AAATGACTAG AGTAGAATGA CAGCTAGCAG ATCTCTTCCC TCCTGCTCCC AGCGCACACA  3360
     AACCCGCCCT CCCCTTGGTG TTGGCGGTCC CTGTGGCCTT CACTTTGTTC ACTACCTGTC  3420
     AGCCCAGCCT GGGTGCACAG TAGCTGCAAC TCCCCATTGG TGCTACCTGG CTCTCCTGTC  3480
     TCTGCAGCTC TACAGGTGAG GCCCAGCAGA GGGAGTAGGG CTCGCCATGT TTCTGGTGAG  3540
     CCAATTTGGC TGATCTTGGG TGTCTGAACA GCTATTGGGT CCACCCCAGT CCCTTTCAGC  3600
     TGCTGCTTAA TGCCCTGCTC TCTCCCTGCC CCACCTTATA GAGAGCCCAA AGAGCTCCTG  3660
     TAAGAGGGAG AACTCTATCT GTGGTTTATA ATCTTGCACG AGGCACCAGA GTCTCCCTGG  3720
     GTCTTGTGAT GAACTACATT TATCCCCTTT CCTGCCCCAA CCACAAACTC TTTCCTTCAA  3780
     AGAGGGCCTG CCTGGCTCCC TCCACCCCAAC TGCACCCATG AGACTCGGTC CAAGAGTCCA  3840
     TTCCCCAGGT GGGAGCCAAC TGTCAGGGAG GTCTTTCCCA CCAAACATCT TTCAGCTGCT  3900
     GGGAGGTGAC CATAGGGCTC TGCTTTTAAA GATATGCTG CTTCAAAGGC CAGAGTCACA  3960
     GGAAGGACTT CTTCCAGGGA GATTAGTGGG GATGGAGAGG AGAGTTAAAA TGACCTCATG  4020
     TCCTTCTTGT CCACGGTTTT GTTGAGTTTT CACTCTTCTA ATGCAAGGGT CTCACACTGT  4080
     GAACCACTTA GGGATGTGATC ACTTTCAGGT GGCCAGGAAT GTTGAATGTC TTTGGCTCAG  4140
     TTCATTTAAA AAAGATATCI ATTTGAAAGT TCTCAGAGTT GTACATATGT TTCACAGTAC  4200
```

```
AGGATCTGTA CATAAAAGTT TCTTTCCTAA ACCATTCACC AAGAGCCAAT ATCTAGGCAT  4260
TTTCTTGGTA GCACAAATTT TCTTATTGCT TAGAAAATTG TCCTCCTTGT TATTTCTGTT  4320
TGTAAGACTT AAGTGAGTTA GGTCTTTAAG GAAAGCAACG CTCCTCTGAA ATGCTTGTCT  4380
TTTTTCTGTT GCCGAAATAG CTGGTCCTTT TTCGGGAGTT AGATGTATAG AGTGTTTGTA  4440
TGTAAACATT TCTTGTAGGC ATCACCATGA ACAAAGATAT ATTTTCTATT TATTTATTAT  4500
ATGTGCACTT CAAGAAGTCA CTGTCAGAGA AATAAAGAAT TGTCTTAAAT GTCATGATTG  4560
GAGATGTCCT TTGCATTGCT TGGAAGGGGT GTACCTAGAG CCAAGGAAAT TGGCTCTGGT  4620
TTGGAAAAAT TTTGCTGTTA TTATAGTAAA CATACAAAGG ATGTCAAAAA AAAAAAAAAA  4680
AAAAAAAAAA AAAAAAAAAA AA
```

Seq ID NO: 142 Protein sequence
Protein Accession #: Eos sequence

```
1          11         21         31         41         51
|          |          |          |          |          |
MGAAGRQDFL FKAMLTISWL TLTCFPGATS TVAAGCPDQS PELQPWNPGH DQDHRVHIGQ    60
GKTLLLTSSA TVYSIHISBG GKLVIKDHDE PIVLRTRHIL IDNGGELHAG SALCPFQGNF   120
TIILYGRADE GIQFDPYYGL KYIGVGKGGA LELHGQKKLS WTFLNKTLHP GGMAEGGYFF   180
ERSWGHRGVI VHVIDPKSGT VIHSDRFDTY RSKKEBERLV QYLNAVPDGR ILSVAVNDEG   240
SBNLDDMARK AMTKLGSKHF LHLGFRHPWS FLTVKGNPSS BVEDHIBYHG HRGSAAARVF   300
KLFQTEHGEY FNVSLSSEWV QDVEWTEWFD BDKVSQTKGG EKISDLWKAH PGKICNRPID   360
IQATTMDGVN LSTEVVYKKG QDYRFACYDR GRACRSYRVR FLCGKPVRPK LTVTIDTNVN   420
STILNLEDNV QSWKPGDTLV IASTDYSMYQ AEEFQVLPCR SCAPNQVKVA GKPMYLHIGE   480
EIDGVDMRAE VGLLSRNIIV MGEMEDKCYP YRNHICNFFD FDTFGGHIKF ALGFKAAHLE   540
GTELKHMGQQ LVGQYPIHPH LAGDVDERGG YDPPTYIRDL SIHHTFSRCV TVHBSNGLLI   600
KDVVGYNSLG HCFFTEDGPE ERNTFDHCLG LLVKSGTLLP SDRDSKMCKM ITEDSYPGYI   660
PKPRQDCNAV STPWMANPNN NLINCAAAGS EETGFWFIPH HVPTGPSVGM YSPGYSEHIP   720
LGKFYNNRAH SNYRAGMIID NGVKTTEASA KDKRPPLSII SARYSPHQDA DPLKPREPAI   780
IRHFIAYKNQ DHGAWLRGGD VWLDSCHFRG EAQEGFLLTG MKAGGILLGG DEAASGMAQG   840
FSPPCRCLLK LVTTGSPFAH VSLAHS
```

Seq ID NO: 143 NM DNA sequence
Nucleic Acid Accession #: NM_006017
Coding sequence: 38..2635

```
1          11         21         31         41         51
|          |          |          |          |          |
CCAAGTTCTA CCTCATGTTT GGAGGATCTT GCTAGCTATG GCCCTCGTAC TCGGCTCCCT    60
GTTGCTGCTG GGGCTGTGCG GGAACTCCTT TTCAGGAGGG CAGCCTTCAT CCACAGATGC   120
TCCTAAGGCT TGGAATTATG AATTGCCTGC AACAAATTAT GAGACCCAAG ACTCCCATAA   180
AGCTGGACCC ATTGGCATTC TCTTTGAACT AGTGCATATC TTTCTCTATG TGGTACAGCC   240
GCGTGATTTC CCAGAAGATA CTTTGAGAAA ATTCTTACAG AAGGCATATG AATCCAAAT   300
TGATTATGAC AAGCCAGAAA CTGTAATCTT AGGTCTAAAG ATTGTCTACT ATGAAGCAGG   360
GATTATTCTA TGCTGTGTCC TGGGGCTGCT GTTTATTATT CTGATGCCTC TGGTGGGGTA   420
TTTCTTTTGT ATGTGTCGTT GCTGTAACAA ATGTGGTGGA GAAATGCACC AGCGACAGAA   480
GGAAAATGGG CCCTTCCTGA GGAAATGCTT TGCAATCTCC CTGTTGGTGA TTTGTATAAT   540
AATAAGCATT GGCATCTTCT ATGGTTTTGT GGCAAATCAC CAGGTAAGAA CCCGGATCAA   600
AAGGAGTCGG AAACTGGCAG ATACGAATTT CAAGGACTCG CGAACTCTCT TGAATGAAAC   660
TCCAGAGCAA ATCAAATATA TATTGGCCCA GTACAACACT ACCAAGGACA AGGCGTTCAC   720
AGATCTGAAC AGTATCAATT CAGTGCTAGG AGGCGGAATT CTTGACCGAC TGAGACCCAA   780
CATCATCCCT GTTCTTGATG AGATTAAGTC CATGGCAACA GCGATCAAGG AGACCAAAGA   840
GGCGTTGGAG AACATGAACA GCACCTTBAA GAGCTTGCAC CAACAAAGTA CACAGCTTAG   900
CAGCAGTCTG ACCAGCGTGA AAACTAGCCT GCGGTCATCT CTCAATGACC CTCTGTGCTT   960
GGGTGCATCCA TCAAGTGAAA CCTGCAACGA CATCAGATTG TCTCTAAGCG AGCTGAATAG  1020
CAACCCTGAA CTGAGGCAGC TTCCACCGGT GGATGCAGAA CTTGACAACG TTAATAACGT  1080
TCTTAGGACA GATTTGGATG GCCTGGTCCA ACAGGGCTAT CAATCCCTTA ATGATATACC  1140
TGACAGAGTA CAACGCCAAA CCACGACTGT CGTAGCAGGT ATCAAAAGGG TCTTGAATTC  1200
CATTGGTTCA GATATCGACA ATGTAACTCA GCGTCTTCCT ATTCAGGATA TACTCTCAGC  1260
ATTCTCTGTT TATGTTAATA ACACTGAAAG TTACATCCAC AGAAATTTAC CTACATTGGA  1320
AGAGTATGAT TCATACTGGT GGCTGGGTGG CCTGGTCATC TGCTCTCTGC TGACCCTCAT  1380
CGTGATTTTT TACTACCTGG GCTTACTGTG TGGCTATGCA GGCATGCCAC GGCATGCAAC  1440
CCCGACCACC CGAGGCTGTG TCTCCAACAC CGGAGGCGTC TTCCTCATGG TTGGAGTTGG  1500
ATTAAGTTTC CTCTTTTTGCT GGATATTGAT GATCATTGTG GTTCTTACCT TTGTCTTTGG  1560
TGCAAATGTG GAAAAACTGA TCTGTGAACC TTACACGAGC AAGGAATTAT TCCGGGTTTT  1620
GGATACACCC TACTTACTAA ATGAAGACTG GGAATACTAT CTCTCTGGGA AGCTATTTAA  1680
TAAATCAAAA ATGAAGCTCA CTTTTGAACA AGTTTACAGT GACTGCAAAA AAAATAGAGG  1740
CACTTACGGC ACTCTTCACC TGCAGAACGA CTTCAATATC AGTGAACATC TCAACATTAA  1800
TGAGCATACT GGAAGCATAA GCAGTGAATT GGAAGTCTG AAGGTAAATC TTAATATCTT  1860
TCTGTTGGGT GCAGCAGGAA GAAAAAACCT TCAGGATTTT GCTGCTTGTG GAATAGACAG  1920
AATGAATTAT GACAGCTACT TGGCTCAGAC TGGTAAATGC CCCGCAGGAG TGAATCTTTT  1980
ATCATTTGCA TATGATCTAG AAGCAAAAGC AAAACAGTTTG CCCCCAGGAA ATTTGAGGAA  2040
CTCCCTGAAA AGAGATGCAC AAACTATTAA AACAATTCAC CAGCAACGAG TCCTTCCTAT  2100
AGAACAATCA CTGAGCACTC TATACCAAAG CGTCAAGACA CTTCAACGCA CAGGGAATGG  2160
ATTGTTGGAG AGAGTAACTA GGATTCTAGC TTCTCTGGAT TTTGCTCAGA ACTTCATCAC  2220
AAACAATACT TCCTCTGTTA TTATTGAGGA AACTAAGAAG TATGGGAGAA CAATAATAGG  2280
ATATTTTGAA CATTATCTGC AGTGGATCGA GTTCTCATC AGTGAGAAAG TGGCATCGTG  2340
CAAACCTGTG GCCACCGCTC TAGATACTGC TGTTGATGTC TTTCTGTGTA GCTACATTAT  2400
CGACCCCTTG AATTTGTTTT GGTTTGGCAT AGGAAAAGCT ACTGTATTTT TACTTCCGGC  2460
TCTAATTTTT GCGGTAAAAC TGGCTAAGTA CTATCGTCGA ATGGATTCGG AGGACGTGTA  2520
CGATGATGTT GAAACTATAC CCATGAAAAA TATGCAAAAT GGTAATAATG GTTATCATAA  2580
AGATCATGTA TATGGTATTC ACAATCCTGT TATGACAAGC CCATCACAAC ATTGATAGCT  2640
GATGTTGAAA CTGCTTGAGC ATCAGGATAC TCAAAGTGGA AAGGATCACA GATTTTTGGT  2700
AGTTTCTGGG TCTACAAGGA CTTTCCAAAT CCAGGAGCAA CGGCCAGTGGC AACGTAGTGA  2760
CTCAGGCGGG CACCAAGGCA ACGGCACCAT TGGTCTCTGG GTAGTGCTTT AAGAATGAAC  2820
```

```
ACAATCACGT TATAGTCCAT GGTCCATCAC TATTCAAGGA TGACTCCCTC CCTTCCTGTC  2880
TATTTTTGTT TTTTACTTTT TTACACTGAG TTTCTATTTA GACACTACAA CATATGGGGT  2940
GTTTGTTCCC ATTGGATGCA TTTCTATCAA AACTCTATCA AATGTGATGG CTAGATTCTA  3000
ACATATTGCC ATGTGTGGAG TGTGCTGAAC ACACACGATT TTACAGGAAA GATGCATTTT  3060
GTGTACAGTA AACGGTGTAT ATACCTTTTG TTACCACAGA GTTTTTTAAA CAAATGAGTA  3120
TTATAGGACT TTCTTCTAAA TGAGCTAAAT AAGTCACCAT TGACTTCTTG GTGCTGTTGA  3180
AAATAATCCA TTTTCACTAA AAGTGTGTGA AACCTACAGC ATATTCTTCA CGCAGAGATT  3240
TTCATCTATT ATACTTTATC AAAGATTGGC CATGTTCCAC TTGGAAATGG CATGCAAAAG  3300
CCATCATAGA GAAACCTGCG TAACTCCATC TGACAAATTC AAAAGAGAGA GAGAGATCTT  3360
GAGAGAGAAA TGCTGTTCGT TCAAAAGTGG AGTTGTTTTA ACAGATGCCA ATTACGGTGT  3420
ACAGTTTAAC AGAGTTTTCT GTTGCATTAG GATAAACATT AATTGGAGTG CAGCTAACAT  3480
GAGTATCATC AGACTAGTAT CAAGTGTTCT AAAATGAAAT ATGAGAAGAT CCTGTCACAA  3540
TTCTTAGATC TGGTGTCCAG CATGGGATGAA ACCTTTGAGT TTGGTCCCTA AATTTGCATG  3600
AAAGCACAAG GTAAATATTC ATTTGCTTCA GGAGTTTCAT GTTGGATCTG TCATTATCAA  3660
AAGTGATCAG CAATGAAGAA CTGGTCGGAC AAAAATTTAAC GTTGATGTAA TGGCAATTCCA  3720
GATGTAGGCA TTCCCCCCAG GTCTTTTCAT GTGCAGATTG CAGTTCTGAT TCATTTGAAT  3780
AAAAAGGAAC TTGG
```

Seq ID NO: 144 NP Protein sequence
Protein Accession #: NP_006008.1

```
    1         11        21        31        41        51
    |         |         |         |         |         |
MALVLGSLLL LGLCGNSFSG GQPSSTDAPK AWNYELPATN YETQDSHKAG PIGILFELVH   60
IFLYVVQPRD PPRDTLRKFL QKAYBSKIDY DKPETVILGL KIVYYBAGII LCCVLGILFI  120
ILMPLVGYFF CMCRCCNKCG GEMHQRQKEN GPFLRKCPAI SLLVICIIIS IGIFYGFVAN  180
HQVRTRIKRS RKLADSNFKD LRTLLLNETPE QIKYILAQYN TTKDKAFIDL NSINSVLGGG  240
ILDRLRFNII PVLDEIKSMA TAIKETKEAL ENMNSTLKSL HQQSTQLSSS LTSVKTSLRS  300
SLNDPLCLVH PSSETCNSIR LSLSQLNSNP ELRQLPFVDA ELDNVNNVLR TDLDGLVQQG  360
YQSLNDIPDR VQRQTTTVVA GIKRVLNSIG SDIDNVTQRL PIQDILSAFS VYVNNTESYI  420
HRNLPTLEEY DSYWWLGGLV ICSLLTLIVI FYYLGLLCGV CGYDRHATPT TRGCVSNTGG  480
VFLMVGVGLS FLFCWILMII VVLTFVPGAN VEKLICEPYT SKELFRVLDT PYLLNEDWEY  540
YLSGKLFNKS KMKLTFEQVY SDCKKNRGTY GTLHLQNSFN ISEHLNINEH TGSISSELES  600
LKVNLNIFLL GAAGRKNLQD FAACGIDRMN YDSYLAQTGK SPAGVNLLSP AYDLEAKANS  660
LPPGNLRNSL KRDAQTIKTI HQQRVLPIEQ SLSTLYQSVK ILQRTGNGLL ERVTRILASL  720
DFAQNFITNN TSSVIIEETK KYGRTIIGYF EHYLQWIEFS ISEKVASCKP VATALDTAVD  780
VFLCSYIIDP LNLFWFGIGK ATVFLLPALI FAVKLAKYYR RMDSEDVYDD VETIPMKNMS  840
NGNNGYHKDH VYGIHNPVMT SPSQH
```

Seq ID NO: 145 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 38..2605

```
    1         11        21        31        41        51
    |         |         |         |         |         |
CCAAGTTCTA CCTCATGTTT GGAGGATCTT GCTAGCTATG GCCCTCGTAC TCGGCTCCCT   60
GTTGCTGCTG GGGCTGTGCG GGAACTCCTT TTCAGGAGGG CAGCCTTCAT CCACAGATGC  120
TCCTAAGGCT TGGAATTATG AATTGCCTGC AACAAATTAT GAGACCCAAG ACTCCCATAA  180
AGCTGGACCC ATTGGCATTC TCTTTGAACT AGTGCATATC TTTCTCTATG TGGTACAGCC  240
GCGTGATTTC CCAGAAGATA CTTTGAGAAA ATTCTTACAG AAGGCATATG AATCCAAAAT  300
TGATTATGAC ATTGTCTACT ATGAAGCAGG GATTATTCTA TGCTGTGTCC TGGGGCTGCT  360
GTTTATTATT CTGATGCCTC TGGTGGGGTA TTTCTTTTGT ATGTGTCGTT GCTGTAACAA  420
ATGTGGTGGA GAAATGCACC AGCGACAGAA GGAAAATGGG CCCTTCCTGA GGAAATGCTT  480
TGCAATCTCC CTGTTGGTGA TTTGTATAAT AATAAGCATT GGCATCTTCT ATGGTTTTGT  540
GGCAAATCAC CAGGTAAGAA CCCGGATCAA AAGGAGTCGG AAACTGGCAG ATAGCAATTT  600
CAAGGACTTG CGAACTCTCT TGAATGAAAC TCCAGAGCAA ATCAAATATA TATTGGCCCA  660
GTACAACACT ACCAAGGACA AGGCGTTCAC AGATCTGAAC AGTATCAATT CAGTGCTAGG  720
AGGCGGAATT CTTGACGGAC TGAGACCCAA CATCATCCCT GTTCTTGATG AGATTAAGTC  780
CATGGCAACA GCGATCAAGG AGACCAAAGA GGCGTTGGAG AACATGAACA GCACCTTGAA  840
GAGCTTGCAC CAACAAAGTA CACAGCTTAG CAGCAGTCTG ACCAGCGTGA AAACTAGCCT  900
GCGGTCATCT CTCAATGACC CTCTGTGCTT GGTGCATCCA TCAAGTGAAA CCTGCAACAG  960
CATCAGATTG TCTCTAAGCC AGCTGAATAG CAACCCTGAA CTGAGGCAGC TTCCACCCGT  1020
GGATGCAGAA CTTGACAACG TTAATAACGT TCTTAGGACA GATTTGGATG GCCTGGTCCA  1080
ACAGGGCTAT CAATCCCTTA ATGATATACC TGACAGAGTA CAACGCCAAA CCACGACTGT  1140
CGTAGCAGGT ATCAAAAGGG TCTTGAATTC CATTGGTTCA GATATCGACA ATGTAACTCA  1200
GCGTCTTCCT ATTCAGGATA TACTCTCAGC ATTCTCTGTT TATGTTAATA ACACTGAAAG  1260
TTACATCCAC AGAAATTTAC CTACATTGGA AGAGTATGAT TCATACTGGT GGCTGGGTGG  1320
CCTGGTCATC TGCTCTCTGC TGACCCTCAT CGTGATTTTT TACTACCTGG GCTTACTGTG  1380
TGGCGTGTGC GGCTATGACA GGCATGCCAC CCCGACCACC CGAGGCTGTG TCTCCAACAC  1440
CGGAGGCGTC TTCCTCATGG TTGGAGTTGG ATTAAGTTTC CTCTTTTGCT GGATATTGAT  1500
GATCATTGTG GTTCTTACCT TTGTGTTTGG TGCAAATGTG GAAAAACTGA TCTGTGAACC  1560
TTACACGAGC AAGGAATTAT TCCGGGTTTT GGATACACCC TACTTACTAA ATGAAGACTG  1620
GGAATACTAT CTCTCTGGGA AGCTATTTAA TAAATCAAAA ATGAAGCTCA CTTTTGAACA  1680
AGTTTACAGT GACTGCAAAA AAAATAGAGG CACTTACGGC ACTCTTCACC TGCAGAACAG  1740
CTTCAATATC AGTGAACATC TCAACATTAA TGAGCATACT GGAAGCATAA GCAGTGAATT  1800
GGAAAGTCTG AAGGTAAATC TTAATATCTT TCTGTTGGGT GCAGCAGGAA GAAAAAACCT  1860
TCAGGATTTT GCTGCTTGTG GAATAGACAG AATGAATTAT GACAGCTACT TGGCTCAGAC  1920
TGGTAAATCC CCCGCAGGAG TGAATCTTTT ATCATTTGCA TATGATCTAG AAGCAAAAGC  1980
AAACAGTTTG CCCCCAGGAA ATTTGAGGAA CTCCCTGAAA AGAGATGCAC AAACTATTAA  2040
AACAATTCAC CAGCAACGAG TCCTTCCTAT AGAACAATCA CTCAGCACTC TATACCAAAG  2100
CGTCAAGATA CTTCAACGCA CAGGGAATGG ATTGTTGGAG AGAGTAACTA GGATTCTAGC  2160
TTCTCTGGAT TTTGCTCAGA ACTTCATCAC AAACAATACT TCCTCTGTTA TTATTGAGGA  2220
AACTAAGAAG TATGGGAGAA CAATAATAGG ATATTTTGAA CATTATCTGC AGTGGATCGA  2280
GTTCTCTATC AGTGAGAAAG TGGCATCGTG CAAACCTGTG GCCACCGCTC TAGATACTGC  2340
```

```
TGTTGATGTC TTTCTGTGTA GCTACATTAT CGACCCCTTG AATTTGTTTT GGTTTGGCAT 2400
AGGAAAAGCT ACTGTATTTT TACTTCCGGC TCTAATTTTT GCGGTAAAAC TGGCTAAGTA 2460
CTATCGTCGA ATGGATTCGG AGGACGTGTA CGATGATGTT GAAACTATAC CCATGAAAAA 2520
TATGGAAAAT GGTAATAATG GTTATCATAA AGATCATGTA TATGGTATTC ACAATCCTGT 2580
TATGACAAGC CCATCACAAC ATTGATAGCT GATGTTGAAA CTGCTTGAGC ATCAGGATAC 2640
TCAAAGTGGA AAGGATCACA GATTTTTGGT AGTTTCTGGG TCTACAAGGA CTTTCCAAAT 2700
CCAGGAGCAA CGCCAGTGCA AACGTAGTGA CTCAGGCGGG CACCAAGGCA ACGGCACCAT 2760
TGGTCTCTGG GTAGTGCTTT AAGAATGAAC ACAATCACGT TATAGTCCAT GGTCCATCAC 2820
TATTCAAGGA TGACTCCCTC CCTTCCTGTC TATTTTTGTT TTTTACTTTT TTACACTGAG 2880
TTTCTATTTA GACACTACAA CATATGGGGT GTTTGTTCCC ATTGGATGCA TTTCTATCAA 2940
AACTCTATCA AATGTGATGG CTAGATTCTA ACATATTGCC ATGTGTGGAG TGTGCTGAAC 3000
ACACACCAGT TTACAGGAAA GATGCATTTT GTGTACAGTA AACGGGTGTAT ATACCTTTTG 3060
TTACCACAGA GTTTTTTAAA CAAATGAGTA TTATAGGACT TTCTTCTAAA TGAGCTAAAT 3120
AAGTCACCAT TGACTTCTTG GTGCTGTTGA AAATAATCCA TTTTCACTAA AAGTGTGTGA 3180
AACCTACAGC ATATTCTTCA CGCAGAGATT TTCATCTATT ATACTTTATC AAAGATTGGC 3240
CATGTTCCAC TTGGAAATGG CATGCAAAAG CCATCATAGA GAAACCTGCG TAACTCCATC 3300
TGACAAATTC AAAAGAGAGA GAGAGATCTT GAGAGAGAAA TGCTGTTCGT TCAAAAGTGG 3360
AGTTGTTTTA ACAGATGCCA ATTACGGTGT ACAGTTTAAC AGAGTTTTCT GTTGCATTAG 3420
GATAAACATT AATTGGAGTG CAGCTAACAT GAGTATCATC AGACTAGTAT CAAGTGTTCT 3480
AAAATGAAAT ATGAGAAGAT CCTGTCACAA TTCTTAGATC TGGTGTCCAG CATGGATGAA 3540
ACCTTTGAGT TTGGTCCCTA AATTTGCATG AAAGCACAAG GTAAATATTC ATTTGCTTCA 3600
GGAGTTTCAT GTTGGATCTG TCATTATCAA AAGTGATCAG CAATGAAGAA CTGGTCGGAC 3660
AAAATTTAAC GTTGATGTAA TGGAATTCCA GATGTAGGCA TTCCCCCCAG GTCTTTTCAT 3720
GTGCAGATTG CAGTTCTGAT TCATTTGAAT AAAAAGGAAC TTGG
```

Seq ID NO: 146 Protein sequence
Protein Accession #:

```
1           11          21          31          41          51
|           |           |           |           |           |
MALVLGSLLL  LGLCGNSFSG  GQPSSTDAPK  AWNYELPATN  YETQDSHKAG  PIGILFELVH   60
IFLYVVQPRD  FPEDTLRKFL  QKAYESKIDY  DIVYYEAGII  LCCVLGLLFI  ILMPLVGYFF  120
CMCRCCNKCG  GEMHQRQKEN  GPFLRKCFAI  SLLVICIIIS  IGIFYGFVAN  HQVRTRIKRS  180
RKLADSNFKD  LRTLLNETPE  QIKYILAQYN  TTKDKAFTDL  NSINSVLGGG  ILDRLRPNII  240
PVLDEIKSMA  TAIKETKEAL  ENMNSTLKSL  HQQSTQLSSS  LTSVKTSLRS  SLNDPLCLVH  300
PSSETCNSIR  LSLSQLNSNP  ELRQLPPVDA  ELDNVNNVLR  TDLDGLVQQG  YQSLNDIPDR  360
VQRQTTTVVA  GIKRVLNSIG  SDIDNVTQRL  PIQDILSAFS  VYVNNTESYI  HRNLPTLEEY  420
DSYWRLGGLV  ICSLLTLIVI  FYYLGLLCGV  CGYDRHATPT  TRGCVSNTGG  VFLMVGVGLS  480
FLFCWILMII  VVLTFVPGAN  VEKLICEPYT  SKELFRVLDT  PYLLNEDWEY  YLSGKLFNKS  540
KMKLTFEQVY  SDCKKNRGTY  GTLHLQNSFN  ISEHLNINEH  TGSISSELES  LKVNLNIFLL  600
GAAGRKNLQD  FAACGIDRMN  YDSYLAQTGK  SPAGVNLLSF  AYDLEAKANS  LPPGNLRNSL  660
KRDAQTIKTI  HQQRVLPIEQ  SLSTLYQSVK  ILQRTGNGLL  ERVTRILASL  DFAQNFITNN  720
TSSVIIEETK  KYGRTIIGYF  EHYLQWIEFS  ISEKVASCKP  VATALDTAVD  VFLCSYIIDP  780
LNLPWFGIGK  ATVFLLDALI  FAVKLAKYYR  RMDSEDVYDD  VETIPMKNME  NGNNGYHKDH  840
VYGIHNPVMT  SPSQH
```

Seq ID NO: 147 DNA sequence
Nucleic Acid Accession #: BC012089.1
Coding sequence: 236..2806

```
1           11          21          31          41          51
|           |           |           |           |           |
ATGCTCTCAG  CTCTCCGCC   GCGGGATGGT  GCCTTGAGTG  AATGACCCCC  TTGGAGAACA   60
TTCTTCGGCA  TCCCTCGCCT  CAAGCCAGCC  TCAGACAGAA  AACTGAAGAT  TCAGCAGATC  120
CAGTGCTTCG  TGCTCCTCTT  CTGCCCAGGA  ACACGCTTGC  CTTCCCCAAG  GCTTCCAGAA  180
GCTCTGAGGC  AGGAGGCACC  AAGTTCTACC  TCATGTTTGG  AGGATCTTGC  TAGCTATGGC  240
CCTCGTACTC  GGCTCCCTGT  TGCTGCTGGG  GCTGTGCGGG  AACTCCTTTT  CAGGAGGGCA  300
GCCTTCATCC  ACAGATGCTC  CTAAGGCTTG  GAATTATGA   TTGCCTGCAA  CAAATTATGA  360
GACCCAAGAC  TCCCATAAAG  CTGGACCCAT  TGGCATTCTC  TTTGAACTAG  TGCATATCTT  420
TCTCTATGTG  GTACAGCCGC  GTGATTTCCC  AGAAGATACT  TTGAGAAAAT  TCTTACAGAA  480
GGCATATGAA  TCCAAAATTG  ATTATGACAA  GATTGTTTAC  TATGAAGCAG  GGATTATTCT  540
ATGCTGTGTC  CTGGGGCTGC  TGTTTATTAT  TCTGATGCCT  CTGGTGGGGT  ATTTTCTTTG  600
TATGTGTCGT  TGCTGTAACA  AATGTGGTGG  AGAAATGCAC  CAGCGACAGA  AGGAAAATGG  660
GCCCTTCCTG  AGGAAATGCT  TTGCAATCTC  CCTGTTGGTG  ATTTGTATAA  TAATAAGCAT  720
TGGCATCTTC  TATGGTTTTG  TGGCAAATCA  CCAGGTAAGA  ACCCGGATCA  AAAGGAGTCG  780
GAAACTGGCA  GATAGCAATT  TCAAGGACTT  GCGAACTCTC  TTGAATGAAA  CTCCAGAGCA  840
AATCAAATAT  ATATTGGCCC  AGTACAACAC  TACCAAGGAC  AAGGCGTTCA  CAGATCTGAA  900
CAGTATCAAT  TCAGTGCTAG  GAGGCGGAAT  TCTTGACCGA  CTGAGACCCA  ACATCATCCC  960
TGTTCTTGAT  GAGATTAAGT  CCATGGCAAC  AGCGATCAAG  GAGACCAAAG  AGGCGTTGGA  1020
GAACATGAAC  AGCACCTTGA  AGAGCTTGCA  CCAACAAAGT  ACACAGCTTA  GCAGCAGTCT  1080
GACCAGCGTG  AAAACTAGCC  TGCGGTCATC  TCTCAATGAC  CCTCTGTGCT  TGGTGCATCC  1140
ATCAAGTGAA  ACCTGCAACA  GCATCAGATT  GTCTCTAAGC  CAGCTGAATA  GCAACCCTGA  1200
ACTGAGGCAG  CTTCCACCCG  TGGATGCAGA  ACTTGACAAC  GTTAATAACG  TTCTTAGGAC  1260
AGATTTGGAT  GGCCTGGTCC  AACAGGGCTA  TCAATCCCTT  AATGATATAC  CTGACAGAGT  1320
ACAACGCCAA  ACCACGACTG  TCGTAGCAGG  TATCAAAAGG  GTCTTGAATT  CCATTGGTTC  1380
AGATATCGAC  AATGTAACTC  AGCGTCTTCC  TATTCAGGAT  ATACTCTCAG  CATTCTCTGT  1440
TTATGTTAAT  AACACTGAAA  GTTACATCCA  CAGAAATTTA  CCTACATTGG  AAGAGTATGA  1500
TTCATACTGG  TGGCTGGGTG  GCCTGGTCAT  CTGCTCTCTG  CTGACCCTCA  TCGTGATTTT  1560
TTACTACCTG  GGCTTACTGT  GTGGCGTGTG  CGGCTATGAC  AGGCATGCCA  CCCCGACCAC  1620
CCGAGGCTGT  GTCTCCAACA  CCGGAGGCGT  CTTCCTCATG  GTTGGAGTTG  GATTAAGTTT  1680
CCTCTTTTGC  TGGATATTGA  TGATCATTGT  GGTTCTTACC  TTTGTCTTTG  GTGCAAATGT  1740
GGAAAAACTG  ATCTGTGAAC  CTTACACGAG  CAAGGAATTA  TTCCGGGTTT  TGGATACACC  1800
CTACTTACTA  AATGAAGACT  GGGAATACTA  CTCTCTCTGGG  AAGCTATTTA  ATAAAATCAA  1860
AATGAAGCTC  ACTTTTGAAC  AAGTTTACAG  TGACTGCAAA  AAAAATAGAG  GCACTTACGG  1920
```

```
CACTCTTCAC CTGCAGAACA GCTTCAATAT CAGTGAACAT CTCAACATTA ATGAGCATAC   1980
TGGAAGCATA AGCAGTGAAT TGGAAAGTCT GAAGGTAAAT CTTAATATCT TTCTGTTGGG   2040
TGCAGCAGGA AGAAAAAACC TTCAGGATTT TGCTGCTTGT GGAATAGACA GAATGAATTA   2100
TGACAGTAC TTGGCTCAGA CTGGTAAATC CCCCGCAGGA GTGAATCTTT TATCATTTGC   2160
ATATGATCTA GAAGCAAAAG CAAACAGTTT GCCCCCAGGA AATTTGAGGA ACTCCCTGAA   2220
AACAGATGCA CAAACTATTA AAACAATTCA CCAGCAACGA GTCCTTCCTA TAGAACAATC   2280
ACTGAGCACT CTATACCAAA GCGTCAAGAT ACTTCAACGC ACAGGGAATG GATTGTTGGA   2340
GAGAGTAACT AGGATTCTAG CTTCTCTGGA TTTTGCTCAG AACTTCATCA CAAACAATAC   2400
TTCCTCTGTT ATTATTGAGG AAACTAAGAA GTATGGGAGA ACAATAATAG GATATTTTGA   2460
ACATTATCTG CAGTGGATCG AGTTCTCTAT CAGTGAGAAA GTGGCATCGT GCAAACCTGT   2520
GGCCACCGCT CTAGATACTG CTGTTGATGT CTTTCTGTGT AGCTACATTA TCGACCCCTT   2580
GAATTTGTTT TGGTTTGGCA TAGGAAAAGC TACTGTATTT TTACTTCCGG CTCTAATTTT   2640
TGCGGTAAAA CTGGCTAAGT ACTATCGTCG AATGGATTCG GAGGACGTGT ACGATGATGT   2700
TGAAACTATA CCCATGAAAA ATATGGAAAA TGGTAATAAT GGTTATCATA AAGATCATGT   2760
ATATGGTATT CACAATCCTG TTATGACAAG CCCATCACAA CATTGATAGC TGATGTTGAA   2820
ACTGCTTGAG CATCAGGATA CTCAAAGTGG AAAGGATCAC AGATTTTTGG TAGTTTCTGG   2880
GTCTACAAGG ACTTTCCAAA TCCAGGAGCA ACGCCAGTGG CAACGTAGTG ACTCAGGCGG   2940
GCACCAAGGC AACGGCACCA TTGGTCTCTG GGTAGTGCTT TAAGAATGAA CACAATCACG   3000
TTATAGTCCA TGGTCCATCA CTATTCTCAAGG ATGACTCCCT CCCTTCCTGT CTATTTTTGT   3060
TTTTTACTTT TTTACACTGA GTTTCTATTT AGACACTACA ACATATGGGG TGTTTGTTCC   3120
CATTGGATGC ATTTCTATCA AAACTCTATC AAATGTGATG GCTAGATTCT AACATATTGC   3180
CATGTGTGGA GTGTGCTGAA CACACACCAG TTTACAGGAA AGATGCATTT TGTGTACAGT   3240
AAACGGTGTA TATACCTTTT GTTACCACAG AGTTTTTTAA ACAAATGAGT ATTATAGGAC   3300
TTTCTTCTAA ATGAGCTAAA TAAGTCACCA TTGACTTCTT GGTGCTGTTG AAAATAATCC   3360
ATTTTCACTA AAAGTGTGTG AAAACCTACAG CATATTCTTC ACGCAGAGAT TTTCATCTAT   3420
TATACTTTAT CAAAGATTGG CCATGTTCCA CTTGGAAATG GCATGCAAAA GCCATCCATAG   3480
AGAAACCTGC GTAACTCCAT CTGACAAATT CAAAAGAGAG ACAGAGATCT TGAGAGAGAA   3540
ATGCTGTTCG TTCAAAGTG GAGTTGTTTT AACAGATGCC AATTACGGTG TACAGTTTAA   3600
CAGAGTTTTC TGTTGCATTA GGATAAACAT TAATTGGAGT GCAGCTAACA TGAGTATCAT   3660
CAGACTAGTA TCAAGTGTTC TAAAATGAAA TATGAGAAGA TCCTGTCACA ATTCTTAGAT   3720
CTGGTGTCCA GCATGGATGA AACCTTTGAG TTTGGTCCCT AAATTTGCAT GAAAGCACAA   3780
GGTAAAATATT CATTTGCTTC AGGAGTTTCA TGTTGGATCT GTCATTATCA AAAGTGATCA   3840
GCAATGAAGA ACTGGTCGGA CAAAATTTAA CGTTGATGTA ATGGAATTCC AGATGTAGGC   3900
ATTCCCCCCA GGTCTTTTCA TGTGCAGATT GCAGTTCTGA TTCATTTGAA TAAAAAGGAA   3960
CTTGGAAAAA AAAAAAAAAA AAA
```

Seq ID NO: 148 Protein sequence
Protein Accession #: AAH12089.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MALVLGSLLL LGLCGNSPSG GQPSSTDAPK AWNYELPATN YETQDSHKAG PIGILFELVH    60
IFLYVVQPRD FPEDTLRKFL QKAYESKIDY DKIVYYEAGI ILCCVLGLLF IILMPLVGYF   120
FCMCRCCNKC GGEMEQRQKE NGPFLRKCPA ISLLVICIII SIGIFYGFVA NEQVRTRIKR   180
SRKLADSNFK DLRTLLNETP EQIKYILAQY NTTKDKAPTD LNSINSVLGG GILDRLRPNI   240
IPVLDEIKSM ATAIKETKEA LENMNSTLKS LHQQSTQLSS SLTSVKTSLR SSLNDPLCLV   300
HPSSETCNSI RLSLSQLNSN PELRQLPPVD AELDNVMNVL RTDLDGLVQQ GYQSLNDIPD   360
RVQRQTTTVV AGIKRVLNSI GSDIDNVTQR LPIQDILSAP SVYVNNTESY IHRNLPTLEE   420
YDSYWWLGGL VICSLLTLIV IFYYLGLLCG VCGYDRHATP TTRGCVSNTG GVFLMVGVGL   480
SFLFCWIIMI IVVLTFVFGA NVEKLICEPY TSKELFRVLD TPYLLNEDWE YYLSGKLFNK   540
SKMKLTFEQV YSDCKKNRGT YGTLHLQNSF NISEHLNINE HTGSISSELE SLKVNLNIFL   600
LGAAGRKNLQ DFARCGIERM NYDSYLAQTG KSPAGVNLLS FAYDLEAKAN SLPPGNLRNS   660
LKRDAQTIKT IHQQRVLPIE QSLSTLYQSV XILQRTGNGL LERVTRILAS LDFAQNFITN   720
NTSSVIIEET KKYGRTIIGY FEHYLQWIEF SISEKVASCK PVATALDTAV DVFLCSYIID   780
PLNLFWFGIG KATVFLLPAL IFAVKLAKYY RRMDSEDVYD DVETIPMKNM ENGNNGYHKD   840
HVYGIHNPVM TSPSQH
```

Seq ID NO: 149 DNA sequence
Nucleic Acid Accession #: NM_033049.1
Coding sequence: 28..1566

```
1          11         21         31         41         51
|          |          |          |          |          |
CCACGCGTCC GAGCAAGAAC AGCTAAAATG AAAGCCATCA TTCATCTTAC TCTTCTTGCT    60
CTCCTTTCTG TAAACACAGC CACCAACCAA GGCAACTCAG CTGATGCTGT AACAACCACA   120
GAAACTGCGA CTAGTGGTCC TACAGTAGCT GCAGCTGATA CCACTGAAAC TAATTTCCCT   180
GAAACTGCTA GCACCACAGC AAATACACCT TCTTTCCCAA CAGCTACTTC ACCTGCTCCC   240
CCCATAATTA GTACACATAG TTCCTCCACA ATTCCTACAC CTGCTCCCCC CATAATTAGT   300
ACACATAGTT CCTCCACAAT TCCTATACCT ACTGCTGCAG ACAGTGAGTC AACTACAAAT   360
GTAAATTCAT TAGCTACCTC TGACATAATC ACCGCTTCAT CTCCAAATGA TGGATTAATC   420
ACAATGGTTC CTTCTGAAAC ACAAAGTAAC AATGAAATGT CCCCCACCAC AGAAGACAAT   480
CAATCATCAG GGCCTCCCAC TGGCACCGCT TTATTGGAGA CCAGCACCCT AAACAGCACA   540
GGTCCCAGCA ATCCTTGCCA AGATGATCCC TGTGCAGATA ATTCGTTATG TGTTAAGCTG   600
CATAATACGA GTTTTTGCCT GTGTTTAGAA GGGTATTACT ACAACTCTTC TACATGTAAG   660
AAAGGAAAGG TATTCCCTGG GAAGATTTCA GTGACAGTAT CAGAAACATT TGACCCAGAA   720
GAGAAACATT CCATGGCCTA TCAAGACTTG CATAGTGAAA TTACTAGCTT GTTTAAAGAT   780
GTATTGGCA CATCTGTTTA TGGACAGACT GTAATTCTTA CTGTAAGCAC ATCTCTGTCA   840
CCAAGATCTG AAATGCGTGC TGATGACAAG TTTGTTAATG TAACAATAGT AACAATTTTG   900
GCAGAAACCA CAAGTGACAA TGAGAAGACT GTGACTGAGA AAATTAATAA AGCAATTAGA   960
AGTAGCTCAA GCAACTTTCT AAACTATGAT TTGACCCTTC GGTGTGATTA TTATGGCTGT  1020
AACCAGACTG CGGATGACTG CCTCTACGGT TTAGCATGCG ATTGCAAATC TGACCTGCAA  1080
AGGCCTACC CACAGAGCCC TTTCTGCGTT GCTTCCAGTC TCAAGTGTCC TGATGCCTGC  1140
AACGCACAGC ACAAGCAATG CTTAATAAAG AAGAGTGGTG GGGCCCCTGA GTGTGCGTGC  1200
GTGCCCGGCT ACCAGGAAGA TGCTAATGGG AACTGCCAAA AGTGTGCATT TGGCTACAGT  1260
```

```
GGACTCGACT GTAAGGACAA ATTTCAGCTG ATCCTCACTA TTGTGGGCAC CATCGCTGGC 1320
ATTGTCATTC TCAGCATGAT AATTGCATTG ATTGTCACAG CAAGATCAAA TAACAAAACG 1380
AAGCATATTG AAGAAGAGAA CTTGATTGAC GAAGACTTTC AAAATCTAAA ACTGCGGTCG 1440
ACAGGCTTCA CCAATCTTGG AGCAGAAGGG AGCGTCTTTC CTAAGGTCAG GATAACGGCC 1500
TCCAGAGACA GCCAGATGCA AAATCCCTAT TCAAGACACA GCAGCATGCC CCGCCCTGAC 1560
TATTAGAATC ATAAGAATGT GGAACCCGCC ATGGCCCCCA ACCAATGTAC AAGCTATTAT 1620
TTAGAGTGTT TAGAAAGACT GATGGGAGAG TGAGCACCAG TAAAGATCTG GCCTCCGGGG 1680
TTTTTCTTCC ATCTGACATC TGCCAGCCTC TCTGAATGGA AGTTGTGAAT GTTTGCAACG 1740
AATCCAGCTC ACTTGCTAAA TAAGAATCTA TGACATTAAA TGTAGTAGAT GCTATTAGCG 1800
CTTGTCAGAG AGGTGGTTTT CTTCAATCAG TACAAAGTAC TGAGACAATG GTTAGGGTTG 1860
TTTTCTTAAT TCTTTTCCTG GTAGGGCAAC AAGAACCATT TCCAATCTAG AGGAAAGCTC 1920
CCCAGCATTG CTTGCTCCTG GGCAAACATT GCTCTTGAGT TAAGTGACCT AATTCCCCTG 1980
GGAGACATAC GCATCAACTG TGGAGGTCCG AGGGGATGAG AAGGGATACC CACCACCTTT 2040
CAAGGGTCAC AAGCTCACTC TCTGACAAGT CAGAATAGGG ACACTGCTTC TATCCCTCCA 2100
ATGGAGAGAT TCTGGCAACC TTTGAACAGC CCAGAGCTTG CAACCTAGCC TCACCCAAGA 2160
AGACTGGAAA GAGACATATC TCTCAGCTTT TTCAGGAGGC GTGCCTGGGA ATCCAGGAAC 2220
TTTTTGATGC TAATTAGAAG GCCTGGACTA AAAATGTCCA CTATGGGGTG CACTCTACAG 2280
TTTTTGAAAT GCTAGGAGGC AGAAGGGGCA GAGAGTAAAA AACATGACCT GGTAGAAGGA 2340
AGAGAGGCAA AGGAAACTGG GTGGGGAGGA TCAATTAGAG AGGAGGCACC TGGGATCCAC 2400
CTTCTTCCTT AGGTCCCCTC CTCCATCAGC AAAGGAGCAC TTCTCTAATC ATGCCCTCCC 2460
GAAGACTGGC TGGGAGAAGG TTTAAAAACA AAAAATCCAG GAGTAAGAGC CTTAGGTCAG 2520
TTTGAAATTG GAGACAAACT GTCTGGCAAA GGGTGCGAGA GGGAGCTTGT GCTCAGGAGT 2580
CCAGCCGTCC AGCCTCGGGG TGTAGGGTTC TGAGGTGTGC CATTGGGGCC TCAGCCTTCT 2640
CTGGTGACAG AGGCTCAGCT GTGGCCACCA CACACACACC ACACACACAC AACCACACAC 2700
ACAAATGGGG GCAACCACAT CCAGTACAAG CTTTTACAAA TGTTATTAGT GTCCTTTTTT 2760
ATTTCTAATG CCTTGTCCTC TTAAAAGTTA TTTTATTTGT TATTATTATT TGTTCTTGAC 2820
TGTTAATTGT GAATGGTAAT GCAATAAAGT GCCTTTGTTA GATGGTGAAA AAAAAAAAAA 2880
AAAAAAA
```

Seq ID NO: 150 Protein sequence
Protein Accession #: NP_149038.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MKAIIHLTLL ALLSVNTATN QGNSADAVTT TETATSGPTV AAADTTETNF PETASTTANT   60
PSFPTATSPA PPIISTHSSS TIPTPAPPII STHSSSTIPI PTAADSESTT NVNSLATSDI  120
ITASSPNDGL ITMVPSETQS NNEMSPTTED NQSSGPPTGT ALLETSTINS TGPSNPCQDD  180
PCADNSLCVK LHNTSFCLCL EGYYYNSSTC KKGKVFPGKI SVTVSETFDP EEKHSMAYQD  240
LHSEITSLFK DVFGTSVYGQ TVILTVSTSL SPRSEMRADD KFVNVTIVTI LAETTSDNEK  300
TVTEKINKAI RSSSSNFLNY DLTLRCDYYG CRQTADDCLN GLACDCKSDL QRPNPQSPFC  360
VASSLKCFDA CNAQHRQCLI KKSGGAPRCA CVPGYQEDAN GNCQKCAFGY SGLDCKDKFQ  420
LILTIVGTIA GIVILSMIIA LIVTARSNNK TKHIEEENLI DEDFQNLKLR STGFTNLGAE  480
GSVFPKVRIT ASRDSQMQNP YSRHSSMPRP DY
```

Seq ID NO: 151 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 87..914

```
1          11         21         31         41         51
|          |          |          |          |          |
AGTACGCGGG GCCGCGTTTT GCCTCCGCAG CAGCTCTGGG CTCTTCTCAG CTGCGCGAGC   60
AGCTGCTCCA ATGCCCCGGA GTGGCCATGG GCGCCCCGCA CTGGTGGGAC CAGCTGCCAGG  120
CTGGTAGCTC GGAGGTGGAC TGGTGCGAGG ACAACTACAC CATCGTGCCT GCTATCGCCG  180
AGTTCTACAA CACGGATCAGC AATGTCTTAT TTTTCATTTT ACCGCCCATC TGCATGTGCT  240
TGTTTCGTCA GTATGCAACA TGCTTCAACA GTGGCATCTA CTTAATCTGG ACTCTTTTGG  300
TTGTAGTGGG AATTGGATCC GTCTACTTCC ATGCAACCCT TAGTTTCTTG GGTCAGATGC  360
TTGATGAACT TGCAGTCCTT TGGGTTCTGA TGTGTGCTTT GGCCATGTGG TTCCCCAGAA  420
GGTATCTACC AAAGATCTTT CGGAATGACC GGGGTAGGTG GTCAGTGTCG  480
TGTCTGCGGT TACGACGTGC CTGGCATTTG TCAAGCCTGC CATCAACAAC ATCTCTCTGA  540
TGACCCTGGG AGTTCCTTGC ACTGCACTGC TCATCGCAGA GCTAAAGAGG TGTGACAACA  600
TGCGTGTGTT TAAGCTGGGC CTCTTCTCGG GCCTCTGGTG GACCCTGGCC CTGTTCTGCT  660
GGATCAGTGA CCGAGCTTTC TGCGAGCTGC TGTCATCCTT CAACTTCCCC TACCTGCACT  720
GCATGTGGCA CATCCTCATC TGCCTTGCTG CCTACCTGGG CTGTGTATGC TTTGCCTACT  780
TTGATGCTGC CTCAGAGATT CCTGAGCAAG GCCCTGTCAT CAAGTTCTGG CCCAATGAGA  840
AATGGGCCTT CATTGGTGTC CCCTATGTGT CCCTCCTGTG TGCCAACAAG AAATCATCAG  900
TCAAGATCAC GTGATGGCAA GATGGTGGCT GGCTTCTCTG CTTATCGCCC CTCATGCAGT  960
GGGCTTCCTT TGCTAGGAAG ACAGCCAAGG GAGTTCGAAT AGTTGGGGTG TGGGCTATCT 1020
TTTCAAAAAT CTATTTGCTG GGGCTCTTAA TTTCTTTAGT GTTCTTTGTA TGTAGGGATT 1080
TAAACTTTGT CATATGGTAC AAATATTCCC TGCCCCCCTG CAGTTCCCA TTTGTCTTTC 1140
AGTATGTTAA TATTTTTGTG CCATACTGGT TTTAAACTTT CATGTTGTCA CATCTGTTAA 1200
TCTTTTCTTT AGGATTTCTG GATTTTGTGT AATTTTTAAA AAGGTCCCCT CCTCCTCCCT 1260
AATGTGTCTG TGGACCACCT GGATTCCACT GTACAAGGGG AAAAGTGTCT ATTCCTTTCC 1320
CAAAGATGGA AAATGGAGGG CTTAGGGACA CTAGATGCAT CTTTCTCAGC ATCACTTCCA 1380
GATGCAGTGA CTTGTTGGGC TGCGTCCTTA ATGGCCATGG CAGAGCAGTC CCTTGGGGGA 1440
TCCCAGCCCTG TACAATGCAT CTCTTCCTGG AGAAAGCTGG CCTGCTCCAG ACCCCACCAT 1500
TCCCAGGCGC CCTTGGAGTG GACTCTACTG ATGACAGACA GACCCTCTGA GAGACAAGAC 1560
CCTCTGACTC TGTGATGGAA GATGCCAGAG ATTTTCCTTT GGGGTAATTG TCCTTAAACA 1620
AAACCAAACA GATGAAACAC ACACAGGACT TGTGGCTAAA AAGGCTAGTT TTTCACTTGC 1680
ATTTCTCAAC TAACCCAGGT TTTACATGCA TCTGTGAATC CTTTTACTAC TACCTCTGTG 1740
GAGAGATGGA GAGACTTCAG ATAAACGTGA AGCTAATGAG TAAAACCCTC TCTGCCAAAA 1800
CCTACACTCC ACTTTAGGCC CTTCTTGAAG ATGAGCACAA TTTTTAAATA CTGAGCACAA 1860
TTTTTAAATA CTGACATCAC TTCCTCTTCC CCCTCCCACC CCAGCTCAGC AGCCTCAAAT 1920
CTACAGAGAA GAAGAATTAT GGCATGAACA TTCCCACAGA CCCACCATCT TTAAGACTTG 1980
ACCTCTGTAA GTTACCAAA GGGCTCCTCA CAATTGTGGT GGGGGTTCTG GTTCAAAATT 2040
```

```
TGGAGCAAAC ATGAAGTTTT TGGAAACGTT TTCTCATTTG AAGCCTCCAG TATGCTGTAC   2100
TATTCTGGAA ATTACCTTCA AGAGTCTCAC TTCTTGTTTC TGTTGTGTTT TCTGTGGGCA   2160
TCATGTTCTT CACGCTTGCA GTAGAAGGTG CTTTCTCGGT TTCCCAGAGT ATCCAACGGC   2220
TCACCTTTCT CAAGTGCTGG CAGTAGCTAT GCACTCACGG GCTGGTTTGG GTCGCTGGTG   2280
CAGCAGCGCA AATCTGTTGC CTTCTGAATT TTTCTCACCT AATGTGACAC TGGCTACAAT   2340
GAATCTTCTC TTCATCGGGC TGAATGAAAG ATTCAAGAAC CATCTTCAAG GTGCATGGTG   2400
GGAATTATCA ACCTCAGGGA TACTCATTTT AACTCAGGCG TGTCCTGCTT TGTAACATTC   2460
CATTGTTGGG AGAGGGCAGG ACAGGTGTGT TCTTCTGTGG GCAGGAGTCA TGTCACTGTC   2520
CTACATATGT AAGAGTTGGG AAGGTGACGA TTTTTGACAC ATCCAGGAAC TCTTACTCTA   2580
GTTAGAATTT GTACCAGATC CAAGGTGAAA ACCCCAATAA GCAACTGAAT TTAGAGTTTA   2640
AAAATGAATG ACTTTATGCT ACATCTGTGG TTATCAAATT ATATAGGTTG TTGAGAAGCA   2700
GAACGCTGTT TGTAGTAAGA AATCTTTGTG GAACCCCAGT GTGTGAAGTA AATTGTATGT   2760
TATTAAATTT ATTTAAGGTT AAATTTTATGG CATTTACTTA ATAATATATG AGGTGGTGAA   2820
AATGCAAATT AACAAATTGG TAATTTCCAA GGTAGAAAAA TTAGGTGTTG AATGAATGTA   2880
TATGTTGGTT TATTATAGTT TTATATATAT ATAGAGAGAG TGTTTTGTTT TTGAGTGCAGG   2940
GTCTTGCTCT ATTACCCAGG CTGGAGTGCA GTGGTGCAGT CATGACTCAC TGTAGCCTCT   3000
GTCTCCCAGG CTCAAGTGAT CCTCTCACCT CAGCCTCCCC AGGAGCTGGG AGTACTGGTG   3060
CGTGCCTCCA CTCCCAGCTA ATTTTTGTAT TTTTTCATAG AGATGGGGTT TCACCATTTT   3120
GCCCAGGCTG GTCTCAAACT TCTGGGCTCA AGAGATTCGC CCGCCTCGGC CTCCCAAAGT   3180
GCTGGGAATA GGCATGAGCC GTCACGCCTG GCCTAAAAAA TATTTTTAAA ATGATCTTTG   3240
AATTAAATAT TCGTAGAATT TCTAATGTAT CTCTTTGAGA CCTAGGAGGT TGATGGAACA   3300
GAACTGCTGT TAAGTCCTTT GGGTTTCAAG TCTAGAATTT TTTAAAGGCA AATATCAGCT   3360
CATTCTTATT TTAGATTGAC CTTATCAGGC ATGGATTCTG GTCTCATCTA CTTTATGGTA   3420
TAAATGCTCC AAGGTAGGGG GTTTGGTATA TATTTTAAGC CCGGCTTTTT TTTTTTTTT   3480
TTTTTTTTTT TTTTTAATGT GAGAAGCAGA ATGTGCTTCT AGAAACTGGT TTTAAAGAGA   3540
TGAGCTGAGA AAGAAATGTG GAATGGAGTA TATTTGAGGA GGACAAAACA TAACTTCACT   3600
TTTGAACAGA AATCACTCTA GCTTGCCAGC ATGGGAATGA AACCAAGAGA GTAGAAATAT   3660
ACCCATCTTA TTTTAAGTTG GGTTTATGGC ATCGCTCATA TATGTAAAAG CACTACAAAC   3720
TCTTTAAAGA AAAATTGGGAA ACTACAGAGA AGTCAAAGAA AAAAAAAAGT AACCCATATT   3780
TCTATTGCCC AGGTATAATC CTTGTTAATA TTTTGGTTTG GTCTCCTCTT TTTTTCCCCH   3840
AATATAGTTG TAAATAAATG ATGTCTTTCA GAGTTGACAT TTATCCTGTA GCTTGAATGG   3900
CATGTAAATG CCAGTTGTAT ATTTTTTCAT GAAGTGTAGG TTTGGAATAC ACTAGAGTTA   3960
GCTATATGCT TGAATGCTGA TCACTGGATT CTGAGACTGA CTACTGAGTC TACCTTTTTA   4020
ATCAAGCCTA ACATGATGGG GCTCCAAAAA GTAATGAATG TAATTGTACT TTTTGATGTG   4080
CCTCTGCACT TGGCTTGGTG AGTCATCATA AATAGCTGTT AAATATGTGA CTTTACAGAT   4140
TTTGATATGT TCAGATTGTA AAAAATGAAT AGTTTATTTC ATTAATTGAT GGGCAGTCAA   4200
GAATCTCCCT CC
```

Seq ID NO: 152 Protein sequence
Protein Accession #: Eos sequence

```
1          11         21         31         41         51
|          |          |          |          |          |
MGAPHWWDQL QAGSSEVDWC EDNYTIVPAI AEFYNTISNV LPFILPPICM CLPRQYATCF    60
NSGIYLIWTL LVVVGIGSVY FHATLSFLGQ MLDELAVLWV LMCALAMWFP RRYLPKIFRN    120
DRGRFKVVVS VLSAVTTCLA FVKPAINNIS LMTLGVPCTA LLIAELXRCD NMRVFKLGLF    180
SGLWWTLALF CWISDRAFCE LLSSFNFPYL HCMWHILICL AAYLGCVCFA YFDAASEIPE    240
QGPVIKFWPN EKWAFIGVPY VSLLCANKKS SVKIT
```

Seq ID NO: 153 DNA sequence
Nucleic Acid Accession #: NM_001432.1
Coding sequence: 167..676

```
1          11         21         31         41         51
|          |          |          |          |          |
TCACTTGCCT GATATTTCCA GTGTCAGAGG GACACAGCCA ACGTGGGGTC CCTTCTAGGC    60
TGACAGCCGC TCTCCAGCCA CTGCCGCGAG CCCGTCTGCT CCCGCCCTGC CCGTGCACTC    120
TCCGCAGCCG CCCTCCGCCA AGCCCCAGCG CCCGCTCCCA TCGCCGATGA CCGCGGGGAG    180
GAGGATGGAG ATGCTCTGTG CCGGCAGGGT CCCTGCGCTG CTGCTCGCC TGGGTTTCCA    240
TCTTCTACAG GCAGTCCTCA GTACAACTGT GATTCCATCA TGTATCCCAG GAGAGTCCGA    300
TGATAACTGC ACAGCTTTAG TTCAGACAGA AGACAATCCA CGTGTGGCTC AAGTGTCAAT    360
AACAAAGTGT AGCTCTGACA TGAATGGCTA TTGTTTGCAT GGACAGTGCA TCTATCTGGT    420
GGACATGAGT CAAAACTACT GCAGGTGTGA AGTGGGTTAT ACTGGTGTCC GATGTGAACA    480
CTTCTTTTTA ACCGTCCACC AACCTTTAAG CAAAGAGTAT GTGGCTTTGA CCGTGATTCT    540
TATTATTTTG TTTCTTATCA CAGTCGTCGG TTCCACATAT TATTTCTGCA GATGGTACAG    600
AAATCGAAAA AGTAAAGAAC CAAAGAAGGA ATATGCAGAG GTTACCTCAG GGGATCCAGA    660
GTTGCCGCAA GTCTGAATGG CGCCATCAAA CTTATGGGCA GGGATAACAG TGTGCCTGGT    720
TAATATTAAT ATTCCATTTT ATTAATAATA TTTATGTTGG GTCAAGTGTT AGGTCAATAA    780
CACTGTATTT TAATGTACTT GAAAAATGTT TTTATTTTTG TTTTATTTTT GACAGACTAT    840
TTGCTAATGT ATAATGTGCA GAAAATATTT AATATCAAAA GAAAATTGAT ATTTTTATAC    900
AAGTAATTTC CTGAGCTAAA TGCTTCATTG AAAGCTTCAA AGTTTATATG CCTGGTGCAC    960
AGTGCTTAGA AGTAAGCAAT TCCCAGGTCA TAGCTCAAGA ATTGTTAGCA AATGACAGAT    1020
TTCTGTAAGC CTATATATAT AGTCAAATCG ATTTTAGTAAG TATGTTTTTT ATGTTCCTCA    1080
AATCAGTGAT AATTGGTTTG ACTGTACCAT GGTTTGATAT GTAGTTGGCA CCATGGTATC    1140
ATATATTAAA ACAATAATGC AATTAGAATT TGGGAGAAGC AAATATAGGT CCTGTGTTAA    1200
ACACTACACA TTTGAAACAA GCTAACCCTG GGGAGTCTAT GGTCTCTTCA CTCAGGTCTC    1260
AGCTATAATT CTGTTATATG AGGGGCAGTG GACAGTTCCC TATGCCAACT CACGACTCCT    1320
ACAGGTACTA GTCACTCATC TACCAGATTC TGCCTATGTA AAATGAATTG AAAAACAATT    1380
TTCTGTAATC TTTTATTTAA GTAGTGGGCA TTTCATAGCT TCACAATGTT CCTTTTTTGT    1440
ATATTACAAC ATTTATGTGA GGTAATTATT GCTCAACAGA CAATTAGAAA AAAGTCCACA    1500
CTGAAGCCT AAATTTGTGC TTTTTAAGAA TATTTTTAGA CTATTTCTTT TTATAGGGGC    1560
TTTGCTGAAT TCTAACATTA AATCACAGCC CAAAATTTGA TGGACTAATT ATTATTTTAA    1620
AATATATGAA GACAATAATT CTACATGTTG TCTTAAGATG GAAATACAGT TATTTCATCT    1680
TTTATTCAAG GAAGTTTTAA CTTTAATACA GCTCAGTAAA TGGCTTCTTC TAGAATGTAA    1740
```

```
AGTTATGTAT TTAAAGTTGT ATCTTGACAC AGGAAATGGG AAAAAACTTA AAAAATTAATA 1800
TGGTGTATTT TTCCAAATGA AAAATCTCAA TTGAAAGCTT TTAAAATGTA GAAACTTAAA 1860
CACACCTTCC TGTGGAGGCT GAGATGAAAA CTAGGGCTCA TTTTCCTGAC ATTTGTTTAT 1920
TTTTTGGAAG AGACAAAGAT TTCTTCTGCA CTCTGAGCCC ATAGGTCTCA GAGAGTTAAT 1980
AGGAGTATTT TTGGGCTATT GCATAAGGAG CCACTGCTGC CACCCACTTTT GGATTTTATG 2040
GGAGGCTCCT TCATCGAATG CTAAACCTTT GAGTAGAGTC TCCCTGGATC ACATACCAGG 2100
TCAGGGAGGA TCTGTTCTTC CTCTACGTTT ATCCTGGCAT GTGCTAGGGT AAACGAAGGC 2160
ATAATAAGCC ATGGCTGACC TCTGGAGCAC CAGGTGCCAG GACTTGTCTC CATGTGTATC 2220
CATGCATTAT ATACCCTGGT GCAATCACAC GACTGTCATC TAAAGTCCTG GCCCTGGCCC 2280
TTACTATTAG GAAAATAAAC AGACAAAAAC AAGTAAATAT ATATGGTCCT ATACATATTG 2340
TATATATATT CATATACAAA CATGTATGTA TACATGACCT TAATGGATCA TAGAATTGCA 2400
GTCATTTGGT GCTCTGCTAA CCATTTATAT AAAACTTAAA AACAAGAGAA AAGAAAAATC 2460
AATTAGATCT AAACAGTTAT TTCTGTTTCC TATTTAATAT AGCTGAAGTC AAAATATGTA 2520
AGAACACATT TTAAATACTC TACTTACAGT TGGCCCTCTG TGGTTAGTTC CACATCTGTG 2580
GATTCAACCA ACCAAGGACG GAAAATGCTT AAAAAATAAT ACAACAACAA CAAAAAATAC 2640
ATTATAACAA CTATTTACTT TTTTTTTTTT CTTTTTGAGA TGGAGTCTCG CTCTGTTGCC 2700
CAGGTTGGAG TGCAGTGGCA CGATCTCGGC TCACTGCAAC CTCACCTCCC GGGTTCAAGA 2760
GATCCTCCTG CCTCAGCCTC CTGAGCAGCT GGGACTACAG GCGCATGCCA CCATGCCCAG 2820
CTAATTTTTG TATTTTTAGT AGAGGCGGGG TTTCACCATG TTGGCCAGGA TGGTCTCAAT 2880
CTCCTAACCT TGAGATCCAC CCTCCCACAGC CTCCCAAACT GCTGGGATTA CAGGCGTGAG 2940
CCACCGCACG TAGCATTTAC ATTAGGTATT ACAAGTAATG TAAAGATGAT TTAAGTATAC 3000
AGGAGGATGT GAATAGGTTA TATGCAAGCA CTATGCCCTT TTATATAAGT GACTTGAACA 3060
TCTGTGCCCG ATTTTAGTAT GTGCAGGGGG GCGATCTGGG AATCAGTCCC CTGTGGATAC 3120
CAAGGTACAA CTGTATTTAT TAACGCTTAC TAGATGTGAG GAGAGTCTGA ATATTTCAG 3180
TGATCTTGGC TGTTTCAAAA AAATCTATTG ACTTTTCAAT AAATCAGCTG CAATCCATTT 3240
ATTTCATTTA CAAAAGATTT ATTGTAAGCC TCTCAATCTT GGTTTTTCAG TTGATCTTAA 3300
GCATGTCAAT TCATAAACAC AAGTCATTTT TGTATTTTTC ATCTTTAAGA ATGCTTAAAA 3360
AAGCTAATCC CTAAAATAGT TAGATCTTTG TAAATGCATA TTAAATAATA AAGTATGACC 3420
CACATTACTT TTTATGGGTG AAAATAAGAC AAAAATAATA GTTTTAGTGA GGATGGTGCT 3480
GAGTAAACAT AAAAACTGAT TTGCTCTCAG CTGATGTGTC CTGTACACAG TGGGAAGATT 3540
TTAGTTCACA CTTAGTCTAA CTCCCCCATT TTACAGATTT CTCACTATAT ATATTTCTAG 3600
AAGGGGCTAT GCATATTCAA TGTATTGAGA ACCAAAGCAA CCACAAATGC ATAAATGCAT 3660
AATTTATGGT CTTCAACCAA GGCCACATAA TAACCCAGTT AACTTACTCT TTAACCAGGA 3720
ATATTAAGTT CTATAACTAG TACTCAAGGT TTAACCTTAA AATTAAGATT TCCTTAACCT 3780
TAACCTTAAA ATTGATATTA TATTAAACAT ACATAATACA ATGTAACTCC ACTGTTCTCC 3840
TGAATATTTT TTGCTCTAAT CTCTCTGCCG AAAGTCAAAG TGATGGGACA ATTGGTATAC 3900
TGGTATGACT ACGTCTTAAG TCAGATTTTT ATTTATGAGT CTTTGAGACT AAATTCAATC 3960
ACCACCAGGT ATCAAATCAA CTTTTATGCA GCAAATATAT GATTCTAGTG TCTGACTTTT 4020
GTTAAATTCA GTAATGCAGT TTTTAAAAAC CTGTATCTGA CCCACTTTGT AATTTTTGCT 4080
CCAATATCCA TTCTGTAGAC TTTTGAAAAA AAAGTTTTTA ATTTGATGCC CAATATATTC 4140
TGACCGTTAA AAAATTCTTG TTCATATGGG AGAAGGGGGA GTAATGACTT GTACAAACAG 4200
TATTTCTGGT GTATATTTTA ATGTTTTTAA AAAGAGTAAT TTCATTTAAA TATCTGTTAT 4260
TCAAATTTGA TGATGTTAAA TGTAAATATAA TGTATTTTCT TTTTATTTTG CACTCTGTAA 4320
TTGCACTTTT TAAGTTTGAA GAGCCATTTT GGTAAACGGT TTTTATTAAA GATGCTATGG 4380
AACATAAAGT TGTATTGCAT GCAATTTAAA GTAACTTATT TGACTATGAA TATTATCGGA 4440
TTACTGAATT GTATCAATTT GTTTGTGTTC AATATCAGCT TTGATAATTG TGTACCTTAA 4500
GATATTGAAG GAGAAAATAG ATAATTTACA AGATATTATT AATTTTTATT TATTTTTCTT 4560
GGGAATTGAA AAAAATTGAA ATAAATAAAA ATGCATTGAA CATCTTGCAT TCAAAATCTT 4620
CACTGAC
```

Seq ID NO: 154 Protein sequence
Protein Accession #: NP_001423.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MTAGRRMEML CAGRVPALLL CLGFHLLQAV LSTTVIPSCI PGESSDNCTA LVQTEDNPRV    60
AQVSITKCSS DMNGYCLHGQ CIYLVDMSQN YCRCEVGYTG VRCEHFFLTV EQPLSKEYVA   120
LTVILIILFL ITVVGSTYYF CRWYRNRKSK EPKKEYERVT SGDPELPQV                
```

Seq ID NO: 155 DNA sequence
Nucleic Acid Accession #: NM_013282.2
Coding sequence: 85..2466

```
1          11         21         31         41         51
|          |          |          |          |          |
CGACTCCTTA GAGCATGGCA TGGCTCAGAG GTGCTGGTAA AACTGATGGG GGTTTTTGCT    60
GTCCCTCCCC TCAGCGCCGA CACCATGTGG ATCCAGGTTC GGACCATGGA CGGGGAGGCAG   120
ACCCACACGG TGGACTCGCT GTCCAGGCTG ACCAAGGTGG AGGACGCTGAG GCGGAAGATC   180
CAGGAGCTGT TCCACGTGGA GCCAGGCCTG CAGAGGCTGT TCTACAGGGG CAAACAGATG   240
GAGGACGGCC ATACCCTCTT CGACTACGAG GTCCGCCTGA ATGACACCAT CCAGCTCCTG   300
GTCCGCCAGA GCCTCGTGCT CCCCCACAGC ACCAAGGAGC GGGACTCCGA GCTCTCCGAC   360
ACCGACTCCG GCTGCTGCCT GGGCCAGAGT GAGTCAGACA AGTCCTCCAC CCACGGCGAG   420
GCGGCCGCCG AGACTGACAG CAGGCCAGCC GATGAGGACA TGTGGGATGA GACGGAATTG   480
GGGCTGTACA AGGTCAATGA GTACGTCGAT GCTCGGGACA CGAACATGGG GGCGTGGTTT   540
GAGGCGCAGG TGGTCAGGGT GACGCGGAAG GCCCCCTCCC GGGACGAGCC CTGCAGCTCC   600
ACGTCCAGGC CGGCGCTGGA GGAGGACGTC ATTTACCACG TGAATACGA CGACTACCCG   660
GAGAACGACG TGGTCCAGAT GAACTCCAGG GACGTCCGAG CGCGGCGCCG CACCATCATC   720
AAGTGGCAGG ACCTGGAGGT GGGCCAGGTG GTCATGCTCA ACTACAACCC CGACAACCCG   780
AAGGAGCGGG GCTTCTGGTA CGACGCGGAG ATCCCAGGA AGCGGCAGAGAC CAGGACGGCG   840
CGGGAACTCT ACGCCAACGT GGTGCTGGGG GATGATTCTC TGAACGACTG TCGGATCATC   900
TTCGTGGACG AAGTCTTCAA GATTGAGCGG CCGGGTGAAG GGAGCCCCAT GGTTGACAAC   960
CCCATGAGAC GGAAGAGCGG GCCGTCCTGC AAGCACTGCA AGGACGACGT GAACAGACTC   1020
TGCCGGGTCT GCGCCTGCCA CCTGTGCGGG GGCCGGCAGG ACCCCGACAA GCAGCTCATG   1080
TGCGATGAGT GCGACATGGC CTTCCACATC TACTGCCTGG ACCCGCCCCT CAGCAGTGTT   1140
```

```
      CCCAGCGAGG  ACGAGTGGTA  CTGCCCTGAG  TGCCGGAATG  ATGCCAGCGA  GGTGGTACTG  1200
      GCGGGAGAGC  GGCTGAGAGA  GAGCAAGAAG  AAGGCGAAGA  TGGCCTCGGC  CACATCGTCC  1260
      TCACAGCGGG  ACTGGGGCAA  GGGCATGGCC  TGTGTGGGCC  GCACCAAGGA  ATGTACCATC  1320
      GTCCCGTCCA  ACCACTACGG  ACCCATCCCG  GGGATCCCCG  TGGGCACCAT  GTGGCGGTTC  1380
      CGAGTCCAGG  TCAGCGAGTC  GGGTGTCCAT  CGGCCCCACG  TGGCTGGCAT  ACACGGCCGG  1440
      AGCAACGACG  GAGCGTACTC  CCTAGTCCTG  GCGGGGGGCT  ATGAGGATGA  CGTGGACCAT  1500
      GGGAATTTTT  TCACATACAC  GGGTAGTGGT  GGTCGAGATC  TTTCCGGCAA  CAAGAGGACC  1560
      GCGGAACAGT  CTTGTGATCA  GAAACTCACC  AACACCAACA  GGGCGCTGGC  TCTCAACTGC  1620
      TTTGCTCCCA  TCAATGACCA  AGAAGGGGCC  GAGGCCAAGG  ACTGGCGGTC  GGGGAAGCCG  1680
      GTCAGGGTGG  TGCGCAATGT  CAAGGGTGGC  AAGAATAGCA  AGTACGCCCC  CGCTGAGGGC  1740
      AACCGCTACG  ATGGCATCTA  CAAGGTTGTG  AAATACTGGC  CGGAGAAGGG  GAAGTCCGGG  1800
      TTTCTCGTGT  GGCGCTACCT  TCTGCGGAGG  GACGATGATG  AGCCTGGCCC  TTGGACGAAG  1860
      GAGGGGAAGG  ACCGGATCAA  GAAGCTGGGG  CTGACCATGC  AGTATCCAGA  AGGCTACCTG  1920
      GAAGCCCTGG  CCAACCGAGA  GCGAGAGAAG  GAGAACCAGCA  AGAGGGAGGA  GGAGGAGCAG  1980
      CAGGAGGGGG  GCTTCGCGTC  CCCCAGGACG  GGCAAGGGCA  AGTGGAAGCG  GAAGTCGGCA  2040
      GGAGGTGGCC  CGAGCAGGGC  CGGGTCCCCG  CGCCGGACAT  CCAAGAAAAC  CAAGGTGGAG  2100
      CCCTACAGTC  TCACGGCCCA  GCAGAGCAGT  CTCATCAGAG  AGGACAAGAG  CAACGCCAAG  2160
      CTGTGGAATG  AGGTCCTGGC  GTCACTCAAG  GACCGGCCGG  CGAGCGGCAG  CCCGTTCCAG  2220
      TTGTTCCTGA  GTAAAGTGGA  GGAGACGTTC  CAGTGTATCT  GCTGTCAGGA  GCTGGTGTTC  2280
      CGGCCCATCA  CGACCGTGTG  CCAGCACAAC  GTGTGCAAGG  ACTGCCTGGA  CAGATCCTTT  2340
      CGGGCACAGG  TGTTCAGCTG  CCCTGCCTGC  CGCTACGACC  TGGCCCGCAG  CTATGCCATG  2400
      CAGGTGAACC  AGCCTCTGCA  GACCGTCCTC  AACCAGCTCT  TCCCCGGCTA  CGGCAATGGC  2460
      CGGTGATCC   CAAGCACTTC  TCGACAGGCG  TTTTGCTGAA  AACGTGTCGG  AGGGCTCGTT  2520
      CATCGGCACT  GATTTTGTTC  TTAGTGGGCT  TAACTTAAAC  AGGTAGTGTT  TCCTCCGTTC  2580
      CCTAAAAAGG  TTTGTCTTCC  TTTTTTTTTA  TTTTTATTTT  TCAAATCTAT  ACATTTTCAG  2640
      GAATTTATGT  ATTCTGGCTA  AAAGTTGGAC  TTCTCAGTAT  TGTGTTAGT   TCTTTGAAAA  2700
      CATAAAAGCC  TGCAATTTCT  CGACAAAACA  ACACAAGATT  TTTTAAAGAT  GGAATCAGAA  2760
      ACTACGTGGT  GTGGAGGCTG  TTGATGTTTC  TGGTGTCAAG  TTCTCAGAAG  TTGCTGCCAC  2820
      CAACTCTTTA  AGAAGGCGAC  AGGATCAGTC  CTTCTCTAGG  GTTCTGGCCC  CCAAGGTCAG  2880
      AGCAAGCATC  TTCCTGACAG  CATTTTGTCA  TCTAAAGTCC  AGTGACATGG  TTCCCCGTGG  2940
      TGGCCCGTGG  CAGCCCGTGG  CATGGCGTGG  CTCAGCTGTC  TGTTGAAGTT  GTTGCAAGGA  3000
      AAAGAGGAAA  CATCTCGGGC  CTAGTTCAAA  CCTTTGCCTC  AAAGCCATCC  CCCACCAGAC  3060
      TGCTTAGCGT  CTGAGATCCG  CGTGAAAAGT  CCTCTCGCCA  GAGTTGGGGC  3120
      CACGCAGAAA  TGGCCTCAAG  GGGACTCTGC  TCCACGTGGG  GCCAGGCGTG  TGACTGACGC  3180
      TGTCCGACGA  AGGCGGCCAC  GGACGGACGC  CAGCACACGA  AGTCACGTGC  AAGTGCCTTT  3240
      GATTCGTTCC  TTCTTTCTAA  AGACGACAGT  CTTTGTTGTT  AGCACTGAAT  TATTGAAAAT  3300
      GTCAACCAGA  TTCTAGAAAC  TGCGGTCATC  CAGTTCTTCC  TGACACCGGA  TGGGTGCTTG  3360
      GGAACCGTTT  GAGCCCTTATA  GATCATTTAC  ATTCAATTTT  TTTAACTCAG  CAAGTGAGAA  3420
      CTTACAAGAG  GGTTTTTTTT  TAATTTTTTT  TTCTCTTAAT  GAACACATTT  TCTAAATGAA  3480
      TTTTTTTTGT  AGTTACTGTA  TATGTACCAA  GAAAGATATA  ACGTTAGGGT  TTGGTTGTTT  3540
      TTGTTTTTGT  ATTTTTTTTC  TTTTGAAAGG  GTTTGTTAAT  TTTTCTAATT  TTACCAAAGT  3600
      TTGCAGCCTA  TACCTCAATA  AAACAGGGAT  ATTTTAAATC  ACATACCTGC  AGACAAACTG  3660
      GAGCAATGTT  ATTTTTTAAAG  GGTTTTTTTC  ACCTCCTTAT  TCTTAGATTA  TTAATGTATT  3720
      AGGGAAGAAT  GAGACAATTT  TGTGTAGGCT  TTTTTCTAAAG  TCCAGTACTT  TGTCCAGATT  3780
      TTAGATTCTC  AGAATAAATG  TTTTTCACAG  ATTGAAAAAA  AAAAAAAA
```

Seq ID NO: 156 Protein sequence
Protein Accession #: NP_037414.2

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      MWIQVRTMDG  RQTETVDSLS  RLTKVEELRR  KIQELFHVEP  GLQRLFYRGK  QMEDGHTLFD  60
      YEVRLNDTIQ  LLVRQSLVLP  HSTKERDSEL  SDTDSGCCLG  QSESDKSSTH  GEAAAETDSR  120
      PADEDMWDET  ELGLYKVNEY  VDARDTNMGA  WPEAQVVRVT  RKAPSRDEPC  SSTSRPALEE  180
      DVIYEVKYDD  YPENGVVQMN  SRDVRARART  IIKWQDLSVG  QVVMLNYNPD  NPKERGFWYD  240
      ABISRKRETR  TARELYANVV  LGDDSLNDCR  IIFVDEVPKI  ERPGEGSPMV  DNPMRRKSGP  300
      SCKHCKDDVN  RLCRVCACHL  CGGRQDPDKQ  LMCDECDMAF  HIYCLDPPLS  SVPSEDEWYC  360
      PECRNDASEV  VLAGERLRES  KKKAKMASAT  SSSQRDWGKG  MACVGRTKEC  TIVPSNHYGP  420
      IPGIPVGTMW  RFRVQVSESG  VHRPHVAGIH  GRSNDGAYSL  VLAGGYEDDV  DKGNFPTYTG  480
      SGGRDLSGNK  RTAEQSCDQK  LTNTNRALAL  NCFAPINDQE  GAEAKDWRSG  KPVRVVRNVK  540
      GGKNSKYAPA  EGNRYDGIYK  VVKYWPEKGK  SGFLVWRYLL  RRDDDEPGPW  TKEGKDRIKK  600
      LGLTMQYPEG  YLEALANRER  EKENSKREEE  EQQEGGFASP  RTGKGKWKRK  SAGGGPSRAG  660
      SPRRTSKKTK  VEPYSLTAQQ  SSLIREDKSN  AKLWNEVLAS  LKDRPASGSP  FQLFLSKVEE  720
      TFQCICCQEL  VFRPITTVCQ  HNVCKDCLDR  SFRAQVFSCP  ACRYDLGRSY  AMQVNQPLQT
```

Seq ID NO: 157 DNA sequence
Nucleic Acid Accession #: NM_000756.1
Coding sequence: 186..776

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      AGAAACTCAG  AGACCAAGTC  CATTGAGAGA  CTGAGGGGAA  AGAGAGGAGA  GAAAGAAAAA  60
      GAGAGTGGGA  ACAGTAAAGA  GAAAGGAAGA  CAACCTCCAG  AGAAAGCCCC  CGGAGACGTC  120
      TCTCTGCAGA  GAGGCGGCAG  CACCCGGCTC  ACCTGCGAAG  CGCCTGGGAA  GCGAGTGCCC  180
      CTAACATGCG  GCTGCCGCTG  CTTGTGTCCG  CGGGAGTCCT  GCTGGTGGCT  CTCCTGCCCT  240
      GCCCGCCATG  CAGGGCCGCTG  CTGAGCCGCG  GGCCGGTCCC  GGGAGCTCGG  CAGGCGCGCGC  300
      AGCACCCTCA  GCCCTTGGAT  TTCTTCCAGC  CGCCGCCGCA  GTCCGAGCAG  CCCCAGCAGC  360
      CGCAGGCTCG  GCCGGTCCTG  CTCCGCATGG  GAGAGGAGTA  CTTCCTCCGC  CTGGGGAACC  420
      TCAACAAGAG  CCCGGCCGCT  CCCCTTTCGC  CCGCCTCCTC  GCTCCTCGCC  GGAGGCAGCG  480
      GCAGCCGGCC  TTCGCCGGAA  CAGGCGACGG  CCAACTTTTT  CCGCGTGTTG  CTGCAGCAGC  540
      TGCTGCTGCC  TCGGCGCTCG  CTCGACAGCC  CCGCGGCTCT  CGCGGAGCGC  GGCGCTAGGA  600
      ATGCCCTCGG  CGGCCACCAG  GAGGGCACCGG  AGAGAGAAAG  GCGGTCCGAG  GAGCCTCCCA  660
      TCTCCCTGGA  TCTCACCTTC  CACCTCCTCC  GGGAAGTCTT  GGAAATGGCC  AGGGCCGAGC  720
      AGTTAGCACA  GCAAGCTCAC  AGCAACAGGA  AACTCATGGA  GATTATTGGG  AAATAAAACG  780
```

```
GTGCGTTTGG CCAAAAAGAA TCTGCATTTA GCACAAAAAA AATTTAAAAA AATACAGTAT    840
TCTGTACCAT AGCGCTGCTC TTATGCCATT TGTTTATTTT TATATAGCTT GAAACATAGA    900
GGGAGAGAGG GAGAGAGCCT ATACCCCTTA CTTAGCACGC ACAAAGTGTA TTCACGTGCA    960
GCAGCAACAC AATGTTATTC GTTTTGTCTA CGTTTAGTTT CCGTTTCCAG GTGTTTATAG   1020
TGGTGTTTTA AAGAGAATGT AGACCTGTGA GAAAACGTTT TGTTTCCAAA AGCAGACAGA   1080
AGTCACTCAA TTGTTTTTGT TGTGGTCTGA GCCAAAGAGA ATGCCATTCT CTTGGGTGGG   1140
TAAGACTAAA TCTGTAAGCT CTTTGAAACA ACTTTCTCTT GTAAACGTTT CAGTAATAAA   1200
ACATCTTTCC AGTCCTTGGT CAGTTTGGTT GTGTAAGAGA ATGTTGAATA CTTATATTTT   1260
TAATAAAAGT TGCAAAGGT
```

Seq ID NO: 158 Protein sequence
Protein Accession #: NP_000747.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MRLPLLVSAG VLLVALLPCP PCRALLSRGP VPGARQAPQH PQPLDFFQPP PQSEQPQQPQ     60
ARPVLLRMGE EYFLRLGNLN KSPAAPLSPA SSLLAGGSGS RPSPEQATAN FFRVLLQQLL    120
LPRRSLDSPA ALAERGARNA LGGHQEAPER ERRSEEPPIS LDLTFHLLRE VLEMARAEQL    180
AQQAHSNRKL MEIIGK
```

Seq ID NO: 159 DNA sequence
Nucleic Acid Accession #: NM_001200.1
Coding sequence: 325..1514

```
1          11         21         31         41         51
|          |          |          |          |          |
GGGGACTTCT TGAACTTGCA GGGAGAATAA CTTGCGCACC CCACTTTGCG CCGGTGCCTT     60
TGCCCCAGCG GAGCCTGCTT CGCCATCTCC GAGCCCCACC GCCCCTCCAC TCCTCGGCCT    120
TGCCCGACAC TGAGCCGCTG TTCCCAGCGT GAAAAGAGAG ACTGCGCGGC CGGCACCCGG    180
GAGAAGGAGG AGGCAAAGAA AAGGAACGGA CATTCGGTCC TTGCGCCAGG TCCTTTGACC    240
AGAGTTTTTC CATGTGGACG CTCTTTCAAT GGACGTGTCC CCGCGTGCTT CTTAGACGGA    300
CTGCGGTCTC CTAAAGGTCG ACCATGGTGG CCGGGACCCG CTGTCTTCTA GCGTTGCTGC    360
TTCCCCAGGT CCTCCTGGGC GGCGCGGGCTG GCCTCGTTCC GGAGCTGGGC CGCAGGAAGT    420
TCGCGGCGGC GTCGTCGGGC CGCCCCTCAT CCCAGCCCTC TGACGAGGTC CTGAGCGAGT    480
TCGAGTTGCG GCTGCTCAGC ATGTTCGGCC TGAAACAGAG ACCCACCCCC AGCAGGGACG    540
CCGTGGTGCC CCCCTACATG CTAGACCTGT ATCGCAGGCA CTCAGGTCAG CCGGGCTCAC    600
CCGCCCCAGA CCACCGGTTG GAGAGGCAG CCAGCCGAGC CAACACTGTG CGCAGCTTCC    660
ACCATGAAGA ATCTTTGGAA GAACTACCAG AAACGAGTGG GAAAACAACC CGGAGATTCT    720
TCTTTAATTT AAGTTCTATC CCCACGGAGG AGTTTATCAC CTCAGCAGAG CTTCAGGTTT    780
TCCGAGAACA GATGCAAGAT GCTTTAGGAA ACAATAGCAG TTTCCATCAC CGAATTAATA    840
TTTATGAAAT CATAAAACCT GCAACAGCCA ACTCGGAAATT CCCCGTGACC AGACTTTTGG    900
ACACCAGGTT GGTGAATCAG AATGCAAGCA GGTGGGAAAG TTTTGATGTC ACCCCCGCTG    960
TGATCGGGTG GACTGCACAG GGACACGCCA ACCATGGATT CGTGGTGGAA GTGGCCCACT   1020
TGGAGGAGAA ACAAGGTGTC TCCAAGAGAC ATGTTAGGAT AAGCAGGTCT TTGCACCAAG   1080
ATGAACACAG CTGGTCACAG ATAAGGCCAT TGCTAGTAAC TTTTGGCCAT GATGGAAAAG   1140
GGCATCCTCT CCACAAAAGA GAAAAACGTC AAGCCAAACA CAAACAGCGG AAACGCCTTA   1200
AGTCCAGCTG TAAGAGACAC CCTTTGTACG TGGACTTCAG TGACGTGGGG TGGAATGACT   1260
GGATTGTGGC TCCCCCGGGG TATCACGCCT TTTACTGCCA CGGAGAATGC CCTTTTCCTC   1320
TGGCTGATCA TCTGAACTCC ACTAATCATG CCATTGTTCA GACGTTGGTC AACTCTGTTA   1380
ACTCTAAGAT TCCTAAGGCA TGCTGTGTCC CGACAGAACT CAGTGCTATC TCGATGCTGT   1440
ACCTTGACGA GAATGAAAAG GTTGTATTAA AGAACTATCA GGACATGGTT GTGGAGGGTT   1500
GTGGGTGTCG CTAGTACAGC AAAATTAAAT ACATAAATAT ATATATA              1500
```

Seq ID NO: 160 Protein sequence
Protein Accession #: NP_001191.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MVAGTRCLLA LLLPQVLLGG AAGLVPELGR RKFAAASSGR PSSQPSDEVL SEFELRLLSM     60
FGLKQRPTPS RDAVVPPYML DLYRRHSGQP GSPAPDHRLE RAASRANTVR SFHHEESLEE    120
LPETSGKTTR RFFFNLSSIP TEEFITSAEL QVFREQMQDA LGNNSSFHHR INIYEIIKPA    180
TANSKFPVTR LLDT
```

Seq ID NO: 161 DNA sequence
Nucleic Acid Accession #: NM_001999.2
Coding sequence: 1..8736

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGGGAGAA GACGGAGGCT GTGTCTCCAG CTCTACTTCC TGTGGCTGGG CTGTGTGGTG     60
CTCTGGGCGC AGGGCACGGC CGGCCAGCCT CAGCCTCCTC CGCCCAAGCC GCCCCGGCCC    120
CAGCCGCCGC CGCAACAGGT TCGGTCCGCT ACAGCAGGCT CTGAAGGCGG GTTTCTAGCG    180
CCCGAGTATC GCGAGGAGGG TGCCGCAGTG GCCAGCCGCG TCCGCCGGCG AGGACAGCAG    240
GACGTGCTCC GAGGGCCCAA CGTGTGCGGC TCCAGATTCC ACTCCTACTG CTGCCCTGGA    300
TGGAAGACGC TCCCTGGAGG AAACCAGTGC ATTGTCCCGA TTTGTAGAAA TAGTTGTGGA    360
GATGGATTTT GTTCCCGTCC TAACATGTGT ACTTGTTCCA GTGGCGAAAT ATCATCAACC    420
TGTGGATCAA AATCAATTCA GCAGTGCAGT GTGAGATGCA TGAATGGTGG GACCTGTGCA    480
GATGACCACT GCCAGTGCCA GAAAGGATAT ATTGGAACTT ATTGTGGACA ACCTGTCTGT    540
GAAAATGGAT GTCAGAATGG TGGACGTGC ATCGCCCAAC CGTGTGCTTG TGTTTATGGG    600
TTCACTGGTC CACAGTGTGA AAGAGATTAC AGGACAGGCC CGTGTTTCAC TCAGGTCAAC    660
AACCAGATGT GCCAAGGGCA GCTGACAGGC ATTGTCTGCA CGAAGACTCT GTGCTGTGCC    720
ACCACTGGAC GGGCGTGGGG CCATCCCTGT GAGATGTGTC CAGCCCAGCC TCAGCCCTGC    780
CGACGGGGTT TCATCCCCAA CATCCGCACT GGGAGCTTGCC AAGATGTTGA TGAATGCCAG    840
GCTATCCCAG GGATATGCCA AGGAGGGAAAC TGTATCAATA CAGTGGGCTC TTTTGAATGC    900
```

```
AGATGCCCTG CTGGTCACAA ACAGAGTGAA ACTACTCAGA AATGTGAAGA CATTGATGAG    960
TGCAGCATCA TTCCTGGGAT ATGTGAAACT GGTGAATGTT CCAACACCGT GGGAAGCTAT   1020
TTTTGTGTTT GTCCACGTGG ATATGTAACC TCAACAGATG GCTCTCGATG CATCGATCAG   1080
AGAACAGGCA TGTGTTTCTC GGGCCTGGTG AATGGCCGCT GTGCACAAGA GCTCCCGGGG   1140
AGAATGACGA AAATGCAGTG CTGCTGTGAG CCTGGCCGCT GCTGGGCAT CGGAACCATT    1200
CCTGAAGCCT GTCCTGTCAG AGGTTCTGAG GAATATCGCA GACTTTGCAT GGATGGACTT   1260
CCAATGGGAG GAATTCCAGG GAGTGCTGGT TCCAGACCTG GAGGCACTGG GGGAAATGGC   1320
TTTGCCCCAA GTGGCAATGG CAATGGCTAT GGCCCAGAGG GGACAGGCTT CATCCCCATC   1380
CCTGGAGGCA ATGGCTTTTC TCCTGGCGTT GGGGGAGCCG GTGTGGGGGC CGGGGGACAG   1440
GGACCTATCA TCACTGGACT AACAATTCTG AACCAGACAA TAGATATCTG TAAGCATCAT   1500
GCTAACCTTT GTTTAAATGG ACGCTGTATA CCAACTGTCT CAAGCTACCG ATGTGAATGC   1560
AACATGGGTT ATAAGCAGGA TGCAAATGGA GATTGTATAG ATGTTGATGA ATGCACATCA   1620
AATCCCTGCA CTAATGGAGA TTGTGTTAAC ACACCTGGTT CCTATTATTG TAAATGTCAT   1680
GCTGGATTCC AGAGGACTCC TACCAAGCAA GCATGCATTG ATATTGATGA GTGCATCCAG   1740
AATGGGGTTC TTTGTAAAAA CGGTCGATGC GTGAACTCAG ATGGAAGTTT CCAGTGCATT   1800
TGCAATGCCG GCTTTGAATT AACTACAGAT GGAAAAAACT GTGTTGATCA TGATGAATGT   1860
ACAACTACCA ACATGTGTTT GAATGGAATG TGCATCAATG AAGATGGCAG CTTCAAGTGC   1920
ATCTGCAAAC CAGGATTTGT CTTGGCTCCA AATGGGCGTT ACTGTACTGA TGTTGATGAA   1980
TGCCAGACCC CAGGAATCTG CATGAATGGG CACTGCATCA ACAGTGAAGG GTCCTTCCGC   2040
TGTGACTGTC CCCCAGGCCT GGCTGTGGGC ATGGATGGAC GTGTGTGTGT TGATACTCAC   2100
ATGCGCAGTA CCTGCTATGG AGGAATCAAG AAAGGAGTGT GTGTGCGTCC TTTCCCCGGT   2160
GCAGTGACCA AGTCCGAATG CTGCTGTGCC AATCCAGACT ATGGTTTTGG AGAACCCTGC   2220
CAGCCATGCC CTGCAAAAAA TTCAGCTGAA TTCCACGGCC TTTGTAGTAG TGGAGTAGGT   2280
ATCACTGTGG ATGGAAGAGA TATCAATGAA TGTGCTTTGA ATCCTGATAT ATGTGCCAAT   2340
GGGATTTGTG AAAACTTACG TGGTAGTTAC CGTTGTAATT GCAACAGTGG CTATGAACCA   2400
GATGCCTCTG GAAGAAACTG TATTGACATT GATGAATGTT TAGTAAACAG ACTGCTTTGT   2460
GATAACGGAT TGTGCCGAAA CACGCCAGGA AGTTACGACT GTACGTGCCC ACCAGGGTAT   2520
GTGTTCAGGA CTGAGACAGA GACCTGTGAA GATATAAATG AATGTGAAAG CAACCCCATGT  2580
GTCAATGGGG CCTGCAGAAA CAACCTTGGA TCTTTCAATT GTGAATGTTC GCCCGGCAGC   2640
AAACTCAGCT CCACAGGATT GATCTGTATT GACAGCCTGA AGGGCACCTG TTGGCTCAAC   2700
ATCCAGGACA GCCGCTGTGA GGTGAATATT AATGGAGCCA CTCTGAAATC TGAATGCTGT   2760
GCCACCCTCG GAGCCGCCTG GGGGAGCCCC TGTGAGCGGT GTGAACTAGA TACAGCTTGC   2820
CCAAGAGGGC TTGCCAGGAT TAAAGGTGTT ACGTGTGAAG ATGTTAATGA GTGTGAGGTG   2880
TTCCCTGGCG TTTGTCCAAA TGGACGCTGT GTCAACAGTA AGGGATCTTT TCATTCGCAG   2940
TGCCCTGAAG GCCTTACGTT GGATGGGACT GGCCGTGTGAT GTTTGGATAT TCGCATGGAG  3000
CAGTGTTACT TGAAGTGGGA TGAAGATGAA TGCATCCACC CCGTTCCTGG AAAGTTCCGC   3060
ATGGATGCCT GCTGCTGTGC TGTCGGGGCB GCTTGGGGCA CCGAGTGTGA GGAGTGCCCC   3120
AAACCTGGCA CCAAGGAATA CGAGCACACTG TGCCCCCGCG GGGCTGGCTT TGCTAACCGA   3180
GGGGATGTTC TTACTGGGCG GCCATTTTAC AAAGACATCA ATGAATGCAA AGCATTTCCT   3240
GGGATGTGCA CTTATGGGAA GTGCAGAAAT ACAATCGGAA GCTTCAAATG CCGTTGCAAT   3300
AGTGGCTTTG CTCTAGACAT GGAGGAAAGA AACTGCACGG ACATCGACGA GTGCAGGATT   3360
TCTCCTGACC TCTGTGGCAG TGGAATCTGC GTCAATACAC CGGGCAGCTT TGAGTGCGAG   3420
TGCTTCGAAG GCTATGAAAG TGGCTTCATG ATGATGAAGA ACTGCATGGA CATTGACGGA   3480
TGTGAACGTA ACCCTCTCCT TTGTAGGGGT GGCACCTGTG TGAACACTGA GGGCAGCTTT   3540
CAGTGTGACT GCCCACTGGG ACACGAGCTG TCACCATCCC GTGAGGACTG TGTGGATATT   3600
AATGAATGCT CCCTGAGTGA CAATCTCTGC AGAAATGGAA AATGTGTGAA CATGATTGGA   3660
ACCTATCAGT GCTCTTGCAA TCCTGGATAT CAGGCTACGC CAGACCGCCA GGGCTGTACA   3720
GATATTGATG AATGTATGAT AATGAACGGA GGCTGTGACA CCCAGTGCAC AAATTCAGAG   3780
GGAAGCTACG AATGCAGCTG CAGTGAGGGT TATGCCCTGA TGCCAGATGG GAGATCGTGT   3840
GCAGACATTG ATGAATGTGA AAACAATCCT GATATCTGTG ATGGCGGCCA GTGTACCAAC   3900
ATTCCTGGAG AGTATCGCTG CCTCTGCTAT GATGGCTTCA TGGCTTCCAT GGACATGAAA   3960
ACATGCATGA ATGTCAATGA ATGTGACCTA AATTCAAATA TCTGCATGTT TGGGGAATGT   4020
GAGAACACAA AGGGATCCTT CATTTGCCAC TGTCAGCTGG GTTACTCAGT GAAGAAGGGG   4080
ACCACAGGAT GTACAGATGT GGATGAGTGT GAAATTGGTG CTCATAACTG CGACATGCAT   4140
GCCTCATGTC TGAATATCCC AGGAAGCTTC AAGTGTAGCT GCAGAGAAGG CTGGATTGGA   4200
AACGGCATCA AGTGTATTGA TCTGGACGAA TGTTCTAATG GAACCCACCA GTGTAGCATC   4260
AATGCTCAGT GTGTAAATAC CCCGGGCTCA TACCGCTGTG CCTGCTCCGA AGGTTTCACT   4320
GGTGATGGCT TTACCTGCTC AGATGTTGAT GAGTGTGCAG AAAACATAAA CCTCTGTGAG   4380
AACGGACAGT GCCTTAATGT CCCGGGTGCA TATCGCTACT AGTGTGAGAT GACCTTCACT   4440
CCCAGCCTCAG ACAGCAGATC CTGCCAAGAT ATTGATGAAT GCTCCTTCCA AAACATTTGT  4500
GTCTCTGGAA CATGTAATAA CCTGCCTGGA ATGTTTCATT GCATCTGCGA TGATGGTTAT   4560
GAATTGGACA GAACAGGAGG GAACTGTACA GATATTGATG AGTGTGCAGA TCCTATAAAC   4620
TGTGTCAATG GCCTATGTGT CAACACGCCT GGTCGCTATG AGTGTAACTG CCCCACCCGAT  4680
TTTCAGTTGA ACCCAACTGG TGTGGGTTGT GTTGACAACC GTGTGGGCAA CTGCTACCTG   4740
AAGTTTGGAC CTCGAGGAGA TGGGAGTCTG TCTTGCAACA CCGAGATCGG GGTGGGCGTC   4800
AGTCGCTCTT CATGCTGCTG CTCTCTGGGA AAGGCCTGGG GAAACCCCTG TGAGACATGC   4860
CCCCCTGTCA ATAGCACTGA ATATTACACC CTGTGTCCCG GAGGTGAAGG CTTCAGACCT   4920
AACCCCATCA CAATCATTTT AGAAGACATT GACGAATGCC AGGAGTTACC AGGTCTCTGC   4980
CAGGGTGGAA ACTGCATCAA CACTTTTGGG AGCTTCCAGT GTGAGTGCCC ACAAGGCTAC   5040
TACCTCAGCG AGGATACCCG CATCTGTGAG GATATTGATG AGTGTTTTGC ACATCCTGGT   5100
GTGTGTGGGC CTGGGACCTG CTATAACACC CTGGGAAATT ACACCTGCAT TTGCCCACCT   5160
GAGTACATGC AGGTCAATGG AGGCCACAAC TGCATGGACA TGAGAAAAAG CTTTTGCTAC   5220
CGAAGCTATA ATGGACCAC TTGTGAGAAT GAGTTGCCTT TCAATGTGAC AAAAAGGATG    5280
TGCTGCTGCA CATATAATGT GGGCAAAGCT GGGAACAAAC CTTGTGAACC ATGCCCAACT   5340
CCAGGAACAG CTGACTTTAA AACCCATATGT GGAAATATTC CTGGATTCAC CTTTGACATT  5400
CACACAGGAA AAGCTGTTGA CATTGATGAA TGTAAAGAGA TTCCAGGCAT TTGTGCAAAT   5460
GGTGTGTGCA TTAACCAGAT TGGCAGTTTC CGCTGTGAAT GCCCTACAGG ATTCAGTTAC   5520
AATGACCTGC TGTTGGTTTG TGAAGATATA GATGAGTGCA GCAATGGTGA TAATCTCTGC   5580
CAGCGGAATG CAGACTGCAT CAATAGTCCT GGTAGTTACC GCTGTGAATG TGCCGCGGGT   5640
TTCAAACTTT CACCCAATGG GGCCTGTGTA GATCGCAATG AATGTTTAGA AATTCCTAAC   5700
GTTTGCAGTC ATGGCTTGTG TGTTGATCTG CAAGGAAGTT ACCAGTGCAT CTGCCACGAT   5760
GGCTTTAAGG CTTCTCAGGA CCAGACCATG TGCATGGATG TTGATGAGTG CGAGCGGCAC   5820
CCATGTGGAA ATGGAACTTG TAAAAACACC GTTGGATCCT ATAACTGTCT GTGCTACCCA   5880
GGGTTTGAAC TCACTCATAA TAATGATTGC CTGGACATAG ATGAGTGCAG TTCCTTTTTT   5940
```

                                    1016

```
GGTCAGGTGT GCAGAAATGG ACGTTGTTTT AATGAAATTG GTTCTTTCAA GTGTCTATGT  6000
AACGAAGGTT ATGAACTTAC CCCAGATGGC AAAAACTGTA TAGACACTAA TGAGTGTGTC  6060
GCCCTTCCCG GCTCTTGCTC TCCTGGTACC TGTCAGAATT TGGAGGGATC CTTCAGATGC  6120
ATCTGTCCCC CAGGGTATGA AGTAAAAAGC GAGAACTGCA TTGATATAAA TGAATGTGAT  6180
GAAGATCCCA ACATTTGTCT TTTTGGTTCC TGTACTAATA CTCCAGGGGG CTTCCAGTGC  6240
CTCTGCCCCC CTGGCTTTGT ACTATCTGAT AATGGACGGA GATGCTTTGA TACTCGCCAG  6300
AGCTTCTGCT TCACAAATTT TGAAAATGGA AAGTGTTCTG TACCCAAAGC TTTCAACACC  6360
ACAAAAGCAA AATGCTGCTG TAGTAAGATG CCAGGAGAGG GCTGGGGGGA CCCCTGTGAG  6420
CTGTGCCCCA AAGACGATGA AGTTGCATTT CAGGATTTGT GTCCATATGG CCATGGAACT  6480
GTCCCTAGTC TTCATGATAC ACGTGAAGAT GTCAATGAGT GTCTTGAGAG CCCAGGCATT  6540
TGTTCAAATG GTCAATGTAT CAACACCGAC GGATCTTTTC GCTGTGAATG TCCAATGGGC  6600
TACAACCTTG ACTACACTGG AGTACGCTGT GTGGATACTG ATGAGTGTTC AATCGGCAAT  6660
CCGTGTGGAA ATGGTACATG CACCAATGTT ATTGGGAGTT TTGAATGCAA TTGCAATGAA  6720
GGCTTTGAGC CAGGGCCCAT GATGAATTGT GAAGATATCA ACGAATGTGC CCAGAACCCA  6780
CTGCTGTGTG CTTTACGCTG CATGAACACT TTTGGGTCCT ATGAATGCAC GTGCCCGATT  6840
GGCTATGCCC TCAGGGAAGA TCAAAGATG TGCAAAGATC TGGATGAATG TGCTGAAGGG  6900
TTACACGACT GTGAATCTAG GGGCATGATG TGTAAGAATC TAATCGGCAC CTTCATGTGC  6960
ATCTGCCCTC CTGGAATGGC CCGAAGGCCC GATGGAGAAG GCTGTGTAGA TGAAAATGAA  7020
TGCAGGACCA AGCCAGGAAT CTGTGAAAAT GGACGTTGTG TTAACATTAT TGGAAGCTAT  7080
AGATGTGAGT GTAATGAAGG ATTCCAGTCA AGTTCTTCAG GCACTGAATG CCTTGACAAT  7140
CGACAGGGTC TCTGCTTTGC AGAGGTACTG CAGACAATAT GTCAAATGGC ATCCAGTAGT  7200
CGCAATCTCG TCACTAAGTC AGAATGCTGC TGTGATGGTG GGCGAGGCTG GGGCCACCAG  7260
TGCGAGCTTT GCCCACTTCC TGGAACTGCC CAGTACAAAA AGATATGTCC TCATGGCCCA  7320
GGATATACAA CTGATGGAAG AGATATTGAT GAATGTAAGG TAATGCCAAA CCTCTGCACC  7380
AATGGTCAGT GCATCAATAC CATGGGCTCA TTCCGATGCT TCTGCAAGGT TGGCTACACC  7440
ACAGACATCA GTGGAACCTC TTGTATAGAC CTTGATGAAT GCTCCCAGTC CCCGAAACCA  7500
TGCAACTACA TCTGCAAGAA CACTGAGGGG AGTTATCAGT GTTCATGTCC GAGGGGGTAT  7560
GTCCTGCAAG AGGATGGAAA GACATGCAAA GACCTTGATG AATGTCAAAC AAAGCAGCAT  7620
AACTGCCAGT TCCTCTGTGT CAACACCCTG GGGGGGTTTA CCTGTAAATG TCCACCTGGT  7680
TTCACACAGC ATCACACTGC TTGTATCGAC AACAACGAAT GTGGGTCTCA ACCTTTGCTT  7740
TGTGGAGGAA AGGGAATCTG TCAAAACACT CCAGGCAGTT TCAGCTGTGA ATGCCAAAGA  7800
GGGGTTCTCTC TTGATGCCAC CGGACTGAAC TGTGAAGATG TTGATGAATG TGATGGGAAC  7860
CACAGGTGCC AACACGGCTG CCAGAACATC CTGGGTGGCT ACAGATGTGG CTGCCCCCAA  7920
GGCTACATCC AGCACTACCA GTGGAATCAG TGTGTCGATG AGAATGAATG CTCCAATCCC  7980
AATGCCTGTG GCTCTGCTTC CTGCTACAAC ACCCTGGGGA GTTACAAGTG CGCCTGCCCC  8040
TCGGGGTTCT CCTTCGACCA GTTCTCCAGT GCCTGCCACG ACGTGAATGA GTGCTCGTCC  8100
TCCAAGAACC CCTGCAATTA CGGCTGCTCT AACACGGAGG GGGGCTACCT CTGTGGCTGC  8160
CCCCCTGGGT ATTACAGAGT GGGACAAGGC CACTGTGTCT CAGGAATGGG ATTTAACAAG  8220
GGGCAGTACC TGTCACTGGA TACAGAGGTC GATGAGGAAA ATGCTCTGTC CCCAGAAGCA  8280
TGCTACGAGT GCAAAATCAA CGGCTATCCT AAGAAGAGACA GCAGGCAGAA GAGAAGTATT  8340
CATGAACCTG ATCCCACTGC TGTTGAACAG ATCAGCCTAG AGAGTGTCGA CATGGACAGC  8400
CCCGTCAACA TGAAGTTCAA CCTCTCCCAC CTCGGCTTCA AGGAGCACAT CCTGGAACTA  8460
AGGCCCGCCA TCCAGCCCCT CAACAACCAC ATCCGGTTATG TCATCTCTCA AGGGAACGAT  8520
GACAGCCGTCT TCCGCATCCA CCAAAGGAAT GGGCTCAGCT ACTTGCACAC GGCCAAGAAG  8580
AAGCTACGAC CCGGCACATA CACACTGGAA ATCACTAGCA TCCCTCTCTA CAAGAAGAAG  8640
GAGCTTAAGA AACTGGAAGA GAGCAATGAG GATGACTACC TCCTAGGGGA GCTTGGGGAG  8700
GCTCTCAGAA TGAGGCTGCA GATTCAGCTC TATTAACCGT TCACAGACTT GGGCCCAGGC  8760
TCAAATCCTA GCACAGCCAG TCTGCAGAAG CATTTGAAAA GTCAAGGACT AATTTTAAAG  8820
AGGAAAAATA ATAATAACTC TTGTTTCTTT CCTCCCTGTC TTAGACTTTG AATGTTGACC  8880
CTCACAGGGA GGGATAATTT AGACTCTGGT ATGGCCAAAG ATTTGAGCTC AAAGGCACCC  8940
GTGGTTACTG TATTTTTTAT ATAACTTCAT TTTAAAATAT ATTAAAAGAA ACCTAAATGT  9000
TCAAGATATC AGCATATGGC ACTAAATGCA CAAAAATAAT GTGAGCTTTT TTTTTTTTT  9060
CCTGTTAGCA GTCTGTAACA CTTTGGGTAT TTTGCTATAG TTGCTAATTA AAAAAATATA  9120
GATGTTTATT TATTTTTAAT GCAGTAATAT ATGGAGAAAT GAACAAACTA TGTAAACAAA  9180
AAGGGAAACT CACTTGTTTT TCTTTAGATT TATAAATTTG AGCTATTTTT TTTAGAGGTG  9240
CTTTTTAAAA ATCCAATAGA TACAAGAGAT GTTTCCTTTG GTTTTCTGCC AGTCATCCAG  9300
CTGATACACA CCTGATCGAT TTTAAAGAAA GCCACACAGA GCTGAATCGG GCAGTGCTAA  9360
TCAATAATTT AAAAGACATG AATGTCATTA GATCCTTTAT AACGTAGATC GAAGCCAAAG  9420
CAGCTCATTT GTGACAACAT TTCATATCAC CAGACACACC AGGCAACAGA AGTTGAAGCA  9480
CAACCACTGT AGCAAAATAC CTTGACTGCT TGTGAGACCA TTAGCATTGC AGGCCAAACC  9540
GTACTGTATT TCCTTCTCAT AACCTCAAGG AACCATATGT GCTACCCACA ACACCTCATT  9600
CTTACCCAGG GTGCGCTGCG TCCTCATGGT ACTGTAGGCA GCTGAAGAAC CGCCGGTTCCC  9660
TTGAAAGGGA ACACCTGGCA TTCTGTGGTG TTTCGTGCTG TCTTAAATAA TGGTGCATTT  9720
ATTATGTTCA AGTTATTTCA GGATTGCCAT ATGTGCAAAC AAATCATGCA ATGCAGCCAA  9780
GGAATATATG TTGTTGTTGT TGTTTAAAAC CCATTTTTTT TTTAGAATTT TCATTAATAC  9840
TGTAGTTATA CACCATATGC CTCATTTTAT CATAGCCTAT TGTGTATGAA AGATGTTTGT  9900
ACAATGAATT GATGTTTAGT TTGCTTTAGT CATTTAAAAA GATATTGTAC CAGGATGTGC  9960
TATTAAGAGC ACGTATCCAT TATTCTTCTC AACCCAAGAA CCTGTTTCCT GGACCAGTGA 10020
CCAAACCTCA TATGTGAAAT GGCCAAAGCA CATGCAGGCT CCTGGTTGTT CCTCTCAAAC 10080
CTGTGCTGAC CAAAGATTAG TAACCAGTTA TACCCAGTAT TTTGAGGTTT TATTGTTTTT 10140
TTAATAACTA AAAAAAAACT CGTTGCC
```

Seq ID NO: 162 Protein sequence
Protein Accession #: NP_001990.1

```
1           11          21        . 31          41          51
I           I           I           I           I           I
MGRRRRLCLQ  LYFLWLGCVV  LWAQGTAGQP  QPPPPKPPRP  QPPPQQVRSA  TAGSEGGFLA   60
PEYREEGAAV  ASEVRRRGQQ  DVLRGPNVCG  SRFHSYCCPG  WKTLPGGNQC  IVPICRNSCG  120
DGFCSRPNMC  TCSSGQISST  CGSKSIQQCS  VRCMNGGTCA  DDHCQCQKGY  IGTYCGQPVC  180
ENGCQNGGRC  IAQPCACVYG  FTGPQCERDY  RTGPCFTQVM  NQMCQGQLTG  IVCTKTLCCA  240
TTGRAWGHPC  EMCPAQPQPC  RRGFIPNIRT  GACQDVDECQ  AIPGICQGGN  CINTVGSFEC  300
RCPAGHKQSE  TTQKCEDIDE  CSIIPGICET  GECSNTVGSY  FCVCPRGYVT  STDGSRCIDQ  360
RTGMCFSGLV  NGRCAQELPG  RMIKMQCCCE  PGRCNGIGTI  PEACPVRGSR  EYRRLCMDGL  420
```

1017

```
        PMGGIPGSAG SRPGGTGGNG PAPSGNGNGY GPGGTGFIPI PGGNGFSPGV GGAGVGAGGQ   480
        GPIITGLTIL NQTIDICKHH ANLCLNGRCI PTVSSYRCEC NMGYKQDANG DCIDVDECTS   540
        NPCTNGDCVN TPGSYYCKCH AGFQRTPTKQ ACIDIDECIQ NGVLCKNGRC VNSDGSFQCI   600
5       CNAGFELTTD GKNCVDHDEC TTTNMCLNGM CINEDGSPKC ICKPGFVLAP NGRYCIDVDE   660
        CQTPGICMNG HCINSEGSFR CDCPPGLAVG MDGRVCVDTH MRSTCYGGIK KGVCVRPFPG   720
        AVTKSECCCA NPDYGPGEPC QPCPAKNSAB FHGLCSSGVG ITVDGRDINE CALDPPDICAN   780
        GICENLRGSY RCNCNSGYEP DASGRNCIDI DECLVNRLLC DNGLCRNTPG SYSCTCPPGY   840
        VPRTETETCE DINECESNPC VNGACRNNLG SFNCBCSPGS KLSSTGLICI DSLKGTCWLN   900
        IQDSRCBVNI NGATLKSECC ATLGAAWGSP CERCELDTAC PRGLARIKGV TCBDVNECEV   960
10      FPGVCPNGRC VNSKGSFHCE CPEGLTLDGT GRVCLDIRME QCYLKWDEDE CIHPVPGKFR   1020
        MDACCCAVGA AWGTECEBCP KPGTKEYETL CPRGAGPANR GDVLTGRPFY KDINECKAPP   1080
        GMCTYGKCRN TIGSFKCRCN SGFALDMBER NCTDIDECRI SPDLCGSGIC VNTPGSPECE   1140
        CPEGYESGFM MMKNCMDIDG CERNPLLCRG GTCVNTEGSF QCDCPLGHEL SPSREDCVDI   1200
15      NECSLSDNLC RNGKCVNMIG TYQCSCNPGY QATPDRQGCT DIDECMIMNG GCDTQCTNSE   1260
        GSYEBCSCSEG YALMPDGRSC ADIDBCESNNP DICDGGQCTN IPGEYRCLCY DGFMASMDMK   1320
        TCIDVNECDL NSNICMFGBC ENTKGSPICH CQLGYSVKKG TTGCTDVDEC EIGAHNCDMH   1380
        ASCLNIPGSF KCSCREGWIG NGIKCIDLDE CSNGTHQCSI NAQCVNTPGS YRCACSEGFT   1440
        GDGPTCSDVD ECAENINLCE NGQCLNVPGA YRCECEMGFT PASDSRSCQD IDECSFQNIC   1500
20      VSGGTCNNLPG MPHCICDDGY ELDRTGGNCT DIDECADPIN CVNGLCVNTP GRYECNCPPD   1560
        FQLNPTGVGC VDNRVGNCYL KFGPRGDGSL SCNTEIGVGV SRSSCCCSLG KAWGNPCETC   1620
        PPVNSTEYYT LCPGGBGFRP NPITIILEDI DECQELPGLC QGGNCINTFG SFQCECPQGY   1680
        YLSBDTRICE DIDECFAHPG VCGPGTCYNT LGNYTCICPP EYMQVNGGHN CMDMRKSFCY   1740
25      RSYNGTTCBN BLPFNVTKRM CCCTYNVGKA GNKPCBPCPT PGTADFKTIC GNIPGPTPDI   1800
        HTGKAVDIDB CKBIPGICAN GVCINQIGSF RCECPTGPSY NDLLLVCEDI DECSNGDNLC   1860
        QRNADCINSP GSYRCBCAAG FKLSPNGACV DRNECLEIPN VCSHGLCVDL QGSYQCICHN   1920
        GFKASQDQTM CMDVDECERH PCGNGTCKNT VGSYNCLCYP GFELTHNNDC LDIDECSBFP   1980
        GQVCRNGRCF NBIGSFKCLC NBGYELTPDG KNCIDTNBCV ALPGSCSPGT CQNLEGSFRC   2040
30      ICPPGYBVKS BNCIDINECD EDPNICLFGS CTNTPGGPQC LCPPGFVLSD NGRRCFDTRQ   2100
        SFCPTNPBNG KCSVPKAPNT TKAKCCCSKM PGEGWGDPCB LCPKDDEVAF QDLCPYGHGT   2160
        VPSLHDTRED VNECLESPGI CSNGQCINTD GSFRCBCPMG YNLDYTGVRC VDTDBCSIGN   2220
        PCGNGTCINV IGSFBCNCNB GPBFGPMMNC BDINECAQNP LLCALRCMNT PGSYECTCPI   2280
35      GYALREDQKM CKDLDECAEG LBDCBSRGMM CKNLIGTFMC ICPPGMARRP DGBGCVDENE   2340
        CRTKPGICEN GRCVNIIGSY RCECNEGFQS SSSGTBCLDN RQGLCFAEVL QTICQMASSS   2400
        RNLVTKSECC CDGGRGNGHQ CELCPLPGTA QYKKICPHGP GYTTDGRDID ECKVMPNLCT   2460
        NGQCINTMGS FRCFCKVGYT TDISGTSCID LDECSQSPKP CNYICKNTEG SYQCSCPRGY   2520
        VLQBDGRTCK DLDBCQTKQH NCQFLCVNTL GGFTCKCPPG FTQHHTACID NNBCGSQFLL   2580
40      CGGKGICQNT PGSFSCBCQR GFSLDATGLN CEDVDBCDGN HRCQHGCQNI LGGYRCGCPQ   2640
        GYIQHYQWNQ CVDENBCSNP NACGSASCYN TLGSYKCACP SGFSFDQFSS ACHDVNBCSS   2700
        SKNPCNYGCS NTEGGYLCGC PPGYYRVGQG HCVSGMGFNK GQYLSLDTBV DEENALSPEA   2760
        CYECKINGYP KKDSRQKRBI HEPDPTAVBQ ISLBSVDMDS PVNMKFNLSH LGSKEHILEL   2820
        RPAIQPLMNH IRYVISQGND DSVFRIHQRN GLSYLHTAKK KLMPGTYTLE ITSIPLYKKK   2880
        BLKKLBBSNB DDYLLGELGE ALRMRLQIQL Y
```

Seq ID NO: 163 DNA sequence
Nucleic Acid Accession #: NM_013372.1
Coding sequence: 63..617

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        GCGGCCGCAC TCAGCGCCAC GCGTCGAAAG CGCAGGCCCC GAGGACCCGC CGCACTGACA   60
        GTATGAGCCG CACAGCCTAC ACGGTGGGAG CCCTGCTTCT CCTCTTGGGG ACCCTGCTGC   120
        CGGCTGCTGA AGGGAAAAAG AAAGGGTCCC AAGGTGCCAA CCCCCCGCCA GACAAGGCCC   180
        AGCACAATGA CTCAGAGCAG ACTCAGTCGC CCCAGCAGCC TGGCTCCAGG AACCGGGGGC   240
        GGGGCCAAGG GCGGGGCACT GCCATGCCCG GGGAGGAGGT GCTGGAGTCC AGCCAAGAGG   300
        CCCTGCATGT GACGGAGCGC AAATACCTGA AGCAGAAACC GTGCAAAACC CAGCCGCTTA   360
        AGCAGACCAT CCACGAGGAA GGCTGCAACA GTCGCACCAT CATCAACCGC TTCTGTTACG   420
        GCCAGTGCAA CTCTTTCTAC ATCCCCAGGC ACATCCGGAA GGAGGAAGGT TCCTTTCAGT   480
        CCTGCTCCTT CTGCAAGCCC AAGAAATTCA CTACCATGAT GGTCACACTC AACTGCCCTG   540
        AACTACAGCC ACCTACCAAG AAGAAGAGAG TCACACGTGT GAAGCAGTGT CGTTGCATAT   600
        CCATCGATTT GGATTAAGCC AAATCCAGGT GCACCCAGCA TGTCCTAGGA ATGCAGCCCC   660
        AGGAAGTCCC AGACCTAAAA CAACCAGATT CTTACTTGGC TTAAACCTAG AGGCCAGAAG   720
        AACCCCCAGC TGCCTCCTGG CAGGAGCCTG CTTGTCGCTA GTTCGTGTGC ATGAGTGTGG   780
        ATGGGTGCCT GTGGGTGTTT TTAGCACCA GAGAAACAC AGTCTCTGCT AGAGAGCACT   840
        CCCTATTTTG TAAACATATC TGCTTCAATG GGGATGTACC AGAAACCCAC CTCACACCCAC   900
        CTCACATCTA AAGGGGCGGG GCCGTGGTCT GGTTCTGACT TTGTGTTTTT GTGCCCCTCCT   960
        GGGGACCAGA ATCTCCTTTC GGAATGAATG TTCATGGAAG AGGCTCCTCT GAGGGCAAGA   1020
        GACCTGTTTT AGTGCTGCAT TCGACATGGA AAAGTCCTTG TAACCTGTGC TTGCATCCTC   1080
        CTTTCCTCCT CCTCCTCACA ATCCATCTCT TCTTAAGTTG ATAGTGACTA TGTCAGTCTA   1140
        ATCTCTTGTT TGCCAAGGTT CCTAAATTAA TTCACTTAAC CATGAATGCAA ATGTTTTTCA   1200
        TTTTGTGAAG ACCCTCCAGA CTCTGGGAGA GGCTGGTGTG GGCAAGGACA AGCAGGATAG   1260
        TGGAGTGAGA AAGGGAGGGT GGAGGGTGAG GCCAAATCGA GTCCAGCAAA AGTCAGTAGG   1320
        GACATTGCAG AAGCTTGAAA GGCCAATACC AGAACACAGG CTGATGCTTC TGAGAAAGTC   1380
        TTTTCCTAGT ATTTAACAGA ACCCAAGTGA ACAGAGGAGA AATGAGGATTG CCAGAAAGTG   1440
        ATTAACTTTG GCCGTTCGCAA TCTGCTCAAA CCTAACACCA AACTCAGAAAC ATAAATACTG   1500
        ACCACTCCTA TGTTCGGACC CAAGCAAGTT AGCTAAACCA AACCAACTCC TCTGCTTTGT   1560
        CCCTCAGGTG GAAAAGAGAG GTAGTTTAGA ACTCTCTGCA TAGGGGTGGG AATTAATCAA   1620
        AAACCCACGA GGCTGAAATT CCTAATACCT TTCCTTTATC GTGGTTATAG TCAGCTCATT   1680
        TCCATTCCAC TATTTCCCAT AATGCTTCTG AGAGCCACTA ACTTGATTGA TAAAGATCCT   1740
        GCCTCTGCTA AGTGTACCTG ACAGTAAGTC TAAAGATGAA AGAGTTTAGG GACTACTCTG   1800
        TTTTAGCCAG ARATATTCTG GGGGTCTTTT TGTTTTAACT ATTGTCAGGA GATTGGGCTA   1860
        RAGAGAAGAC GACGACAGTA AGGAAATAAA GGGRATTGCC TCTGGCTAGA GAGTAAGTTA   1920
        GGTGTTAATA CCTGGTAGAA ATGTAAGGGA TATGACCTCC CTTTCTTTAT GTGCTCACTG   1980
        AGGATCTGAG GGGACCCTGT TAGGAGAGCA TAGCCATCATG ATGTATTAGC TGTTCATCTG   2040
        CTACTGGTTG GATGGACATA ACTATTGTAA CTATTCAGTA TTTACTGGTA GGCACTGTCC   2100
```

```
TCTGATTAAA CTTGGCCTAC TGGCAATGGC TACTTAGGAT TGATCTAAGG GCCAAAGTGC   2160
AGGGTGGGTG AACTTTATTG TACTTTGGAT TTGGTTAACC TGTTTTCTTC AAGCCTGAGG   2220
TTTTATATAC AAACTCCCTG AATACTCTTT TTGCCTTGTA TCTTCTCAGC CTCCTAGCCA   2280
AGTCCTATGT AATATGGAAA ACAAACACTG CAGACTTGAG ATTCAGTTGC CGATCAAGGC   2340
TCTGCCATTC AGAGAACCCT TGCAACTCGA GAAGCTGTTT TTATTTCGTT TTTGTTTTGA   2400
TCCAGTGCTC TCCCATCTAA CAACTAAACA GGAGCCATTT CAAGGCGGGA GATATTTTAA   2460
ACACCCAAAA TGTTGGGTCT GATTTTCAAA CTTTTAAACT CACTACTGAT GATTCTCACG   2520
CTAGGCGAAT TTGTCCAAAC ACATAGTGTG TGTGTTTTGT ATACACTGTA TGACCCCACC   2580
CCAAATCTTT GTATTGTCCA CATTCTCCAA CAATAAAGCA CAGAGTGGAT TTAATTAAGC   2640
ACACAAATGC TAAGGCAGAA TTTTGAGGGT GGGAGAGAAA AAAAGGGAAA GAAGCTGAAA   2700
ATGTAAAACC ACACCAGGGA GGAAAAAATGA CATTCAGAAC AGCAAACAC TGAATTTCTC    2760
TTGTTGTTTT AACTCTGCCA CAAGAATGCA ATTTCGTTAA TGGAGATGAC TTAAGTTGGC   2820
AGCAGTAATC TTCTTTTAGG AGCTTGTACC ACAGTCTTGC ACATAAGTGC AGATTTGGCT   2880
CAAGTAAAGA GAATTTCCTC AACACTAACT TCACTGGGAT AATCAGCAGC GTAACTACCC   2940
TAAAAGCATA TCACTAGCCA AAGAGGGAAA TATCTGTTCT TCTTACTGTG CCTATATTAA   3000
GACTAGTACA AATGTGGTGT GTCTTCCAAC TTTCATTGAA AATGCCATAT CTATACCATA   3060
TTTTATTCGA GTCACTGATG ATGTAATGAT ATATTTTTTC ATTATTATAG TAGAATATTT   3120
TTATGGCAAG ATATTTGTGG TCTTGATCAT ACCTATTAAA ATAATGCCAA ACACCAAATA   3180
TGAATTTTAT GATGTACACT TTGTGCTTGG CATTAAAAGA AAAAAACACA CATCCTGGAA   3240
GTCTGTAAGT TGTTTTTTGT TACTGTAGGT CTTCAAAGTT AAGAGTGTAA GTGAAAATC    3300
TGGAGGAGAG GATAAATTCC ACTGTGTGGA ATGTGAATAG TTAAATGAAA AGTTATGGTT   3360
ATTTAATGTA ATTATTACTT CAAATCCTTT GGTCACTGTG ATTTCAAGCA TGTTTTCTTT   3420
TTCTCCTTTA TATGACTTTC TCTGAGTTGG GCAAAGAAGA AGCTGACACA CCGTATGTTG   3480
TTAGAGTCTT TTATCTGTCC AGGGGAAACA AAATCTTGAC CCAGCTGAAC ATGTCTTCCT   3540
GAGTCAGTGC CTGAATCTTT ATTTTTTAAA TTGAATGTTC CTTAAAGGTT AACATTTCTA   3600
AAGCAATATT AAGAAAGACT TTAAATGTTA TTTTGGAAGA CTTACGATGC ATGTATACAA   3660
ACGAATAGCA GATAATGATG ACTAGTTCAC ACATAAAGTC CTTTTAAGGA GAAAATCTAA   3720
AATGAAAAGT GGATAAACAG AACATTTATA AGTGATCAGT TAATGCCTAA GAGTGAAAGT   3780
AGTTCTATTG ACATTCCTCA AGATATTTAA TATCAACTGC ATTATGTATT ATGTCTGCTT   3840
AAATCATTTA AAAACGGCAA AGAATTATAT AGACTATGAG GTACCTTGCT GTGTAGGAGG   3900
ATGAAAGGGG AGTTGATAGT CTCATAAAAC TAATTTGGCT TCAAGTTTCA TGAATCTGTA   3960
ACTAGAATTT AATTTTCACC CCAATAATGT TCTATATAGC CTTTGCTAAA GAGCAACTAA   4020
TAAATTAAAC CTATTCTTTC AAAAAAAAAA
```

Seq ID NO: 164 Protein sequence
Protein Accession #: NP_037504.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MSRTAYTVGA LLLLLGTLLP AAEGKKKGSQ GAIPPPDKAQ HNDSEQTQSP QQPGSRNRGR   60
GQGRGTAMPG EEVLESSQEA LHVTERKYLK RDWCKTQPLK QTIHEEGCNS RTIINRFCYG   120
QCNSFYIPRH IRKEEGSFQS CSFCKPKKFT TMMVTLNCPE LQPPTKKKRV TRVKQCRCIS   180
IDLD
```

Seq ID NO: 165 DNA sequence
Nucleic Acid Accession #: CAT cluster

```
1          11         21         31         41         51
|          |          |          |          |          |
GAATTGCATC GGACAGAGCT TCGCCATGGC CGAGTTGCGC GTGGTTGTGG CACTAACACT   60
GCTACGTTTC CGCCTGAGCG TGGACCGAAC GCGCAAGGTG CGGCGGAAGC CGGAGCTCAT   120
ACTGCGCACG GAGAACGGGC TCTGGCTCAA G
```

Seq ID NO: 166 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..1650

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCCACCTC TGCCATCCAG AATGTCTCGG GGGCCCTCCT TGGTGACAGG CAGGATGCTG   60
CCCATCACAG ACCGCCTGCT GCACCTCCTG GGGCTGGAGA AGACGGCGTT CCGCATATAC   120
GCGGTGTCCA CCCTTCTCCT CTTCCTGCTC TTCTTCCTGT TCCGCCTGCT GCTGCGGTTC   180
CTGAGGCTCT GCAGGAGCTT CTACATACC TGCCGCCGGC TGCGCTGCTT CCCCCAGCCT    240
CCCCGGCGCA ACTGGCTGCT GGGCCACCTG GGCATGTACC TTCCAAATGA GGCGGGGCCTT  300
CAAGATGAGA AGAAGGTACT GGACAACATG CACCATGTAC TCTTGGTATG GATGGGGACCT  360
GTCCTGCCGC TGTTGGTTCT GGTGCACCCT GATTACATCA AACCCCTTTT GGGAGCCTCA   420
GCTGCCATCG CCCCCAAGGA TGACCTCTTC TATGGCTTCC TAAAACCTTG GCTAGGGGAT   480
GGGCTGCTGC TCAGCAAAGG TGACAAGTGG AGCCGGCACC GTCGCCTGCT GACACCCGCC   540
TTCCACTTTG ACATCCTGAA GCCTTACATG AAGATCTTCA ACCAGAGCGC TGACATTATG   600
CATGCTAAAT GGCGGCATCG GGCAGAGGGC TCAGCGGTCT CCCTTGATAT GTTTGAGCAT   660
ATCAGCCTCA TGACCCTGGA CAGTCTTCAG AAATGTGTCT TCAGCTACAA CAGCAACTGC   720
CAAGAGAAGA TGAGTGATTA TATCTCCGCT ATCATTGAAC TGAGCGCTCT GTCTGTCCGG   780
CGCCAGTATC GCTTGCCACCA CTACCTCGAC TTCATTTACT ACCGCTCGGC GGATGGGCGG   840
AGGTTCCGGC AGGCCTGTGA CATGGTGCAC CACTTCACCA CTGAAGTCAT CCAGGAACGG   900
CGGCGGCCAC TGCGTCAGCA GGGGGCCGAG GCCTGGCTTA AGGCCAAGCA GGGGAAGACC   960
TTGGACTTTA TTGATGTGCT GCTCCTGGCC AGGGATGAAG ATGGAAAGGA ACTGTCAGAC   1020
GAGGATATCC GAGCCGAAGC AGACACCTTC ATGTTTGAGG GTCACGACAC AACATCCAGT   1080
GGGATCTCTT GGATGCTGTT CAATTTGGCA AAGTATCCGG AATACCAGGA GAAATGCCGA   1140
GAAGAGATTC AGGAAGTCAT GAAAGGCCGG GAGCTGGAGTG AGCTGGAGTG GGACGATCTG   1200
ACTCAGCTGC CCTTTACAAC TATGTGCATT AAGGAGAGCC TGCGCCAGTA CCCACCTGTC   1260
ACTCTTGTCT CTCGCCAATG CACGGAGGAC ATCAAGCTCC AGATGGGCG CATCATCCCC   1320
AAAGGAATCA TCTGCTTGGT CAGCATCTAT GGAACCCACC ACAACCCCAC AGTGTGGCCT   1380
GACTCCAAGG TGTACAACCC CTACCGCTTT GACCGGACA ACCCACAGCA GCGCTCTCCA    1440
```

```
CTGGCCTATG TGCCCTTCTC TGCAGGACCC AGGAATTGCA TCGGACAGAG CTTCGCCATG    1500
GCCGAGTTGC GCGTGGTTGT GGCACTAACA CTGCTACGTT TCCGCCTGAG CGTGGCACCGA   1560
ACGCGCAAGG TGCGGCGGAA GCCGGGAGCTC ATACTGCGCA CGGAGAACGG GCTCTGGCTC   1620
AAGGTGGAGC CGCTGCCTCC GCGGGCCTGA
```

Seq ID NO: 167 Protein sequence
Protein Accession #:   FGENESH predicted

```
1          11         21         31         41         51
|          |          |          |          |          |
MPPLPSRMSR GPSLVTGRML PITDRLLHLL GLEKTAFRIY AVSTLLLFLL FFLFRLLLRF    60
LRLCRSFYIT CRRLRCFPQP PRRNWLLGHL GMYLPNEAGL QDEKKVLDNM HHVLLVWMGP    120
VLPLLVLVHP DYIKPLLGAS AAIAPKDDLF YGPLKPWLGD GLLLSKGDKW SRHRRLLTPA    180
PHFDILKPYM KIFNQSADIM HAKWRHLAEG SAVSLDMFEH ISLMTLDSLQ KCVFSYNSNC    240
QEKMSDYISA IIELSALSVR RQYRLHHYLD FIYYRSADGR RFRQACDMVH HFTTEVIQER    300
RRALRQQGAE AWLKAKQGKT LDFIDVLLLA RDEDGKELSD EDIRAEADTF MPEGHDTTSS    360
GISWMLFNLA KYPEYQEKCR EEIQEVMKGR ELEELEWDDL TQLPFTTMCI KESLRQYPPV    420
TLVSRQCTED IKLPDGRIIP KGIICLVSIY GTHHNPTVWP DSKVYNPYRF DPDNPQQRSP    480
LAYVPFSAGP RNCIGQSFAM AELRVVVALT LLRFRLSVDR TRKVRRKPEL ILRTENGLWL    540
KVEPLPPRA
```

Seq ID NO: 168 DNA sequence
Nucleic Acid Accession #: AK058088.1
Coding sequence: 252..1772

```
1          11         21         31         41         51
|          |          |          |          |          |
AGGAAACCAA GGCAAGCTCC CCCTGTCAAA GCACCTTGGC CCATAAGAAG AAAAGGGGGA    60
GCCCCAGATG TGATGAGCGC TTCCAGGCTT CAGGCTCAGA AGGCGCCCCC AGCTCTCCTG    120
TAACTCAGAG GCCAGTGTGA TGGGAGTTCC TCCACTCAGC ACACTTCCCC TGTAAACACG    180
CCTGTGGTGG GCAAAAGGGC TTTGGAACGG TTGCTTGTCT TTTCTCTCCT GCGTAATTTC    240
CACTTTCATT CATGATAATG TCGAACACGC ACAAAGCTCG GCTGGAACGC CGGGTCACTG    300
GCTCAACCAA CCGGTGGCGT TTGCCCAAAC AGCCTTTCTC TGGGGACCTG CTCTCACTTT    360
CCCAGATGTG CAAGGCTCTG AGCATAGACT TTGAGGAAGC TTTGAGGAAC CCAGACAGGT    420
TATGCATTTC ACAAATCCAG AAGTTTTTCT TTGAGAATTT CAAGAACAAG GACATCCAAA    480
GTGGGGAAGC AGATGTGATT CTCGAGTGCC TGGGCTTCAA ATGGGAGCTC CATCAGCCGC    540
GGCTTTTTCA GTCTGAGACC TTGGCCAAGC TCTACCTGAA AGCCCTGGCG CAGGGCACCA    600
CACACCCCCT GAGGGAGCTG GAGGAGCTTC TGCGAGCTCA ATCACCTAAG AAGACCAAAG    660
AAAAATCCCC TGCAAAGAGG ATCATCATTT CCTTGAAGAT CAATGACCCA CTGGTCACTA    720
AAGTCGCCTT CGCCACGGCC CTGAAGAACC TCTACATGAG TGAGGTGGAG ATTAACTTGG    780
AAGACCTACT GGGAGTGCTG GCTTCCGCCC ACATCCTCCA GTTCAGTGGC CTGTTTCAAA    840
GGTGCGTGGA TGTGATGATA GCCAGACTCA AGCCAAGCAC CATCAAGAAA TTCTACGAGG    900
CCGGCTGCAA GTACAAGGAA GAGCAGCTCA CCACCGGCTG CGAGAAGTGG CTGGAAATGA    960
ACTTGGTTCC TCTAGGGGGG ACGCAGATCC ACCTCCACAA AATCCCACAG GACCTGCTCC    1020
ACAAAGTGCT GAAGTCCCCC AGGTTATTTA CCTTTAGTGA ATTCCATCTT CTGAAAACAA    1080
TGCTTTTGTG GGTCTTCTTG CAACTGAACT ACAAGATTCA GGCAATTCCG ACTTATGAAA    1140
CCGTGATGAC ATTTTTTAAG AGCTTTCCTG AGAACTGTTG CTTTCTGGAC CGGGACATAG    1200
GACGGAGCTT GAGGCCGGCTC TTCCTCTGCT TGCGTCTGCA CGGCATCACC AAAGGCAAGG    1260
ATCTGGAGGT GCTGCGGCAC CTTAACTTCT TCCCAGAGTC ATGGCTGGAC CAGGTTACGG    1320
TCAACCATTA CCACGCACTG GAGAATGGGG GCGACATGGT CCACCTGAAA GATCTTAACA    1380
CCCAGCTGT GAGATTTGGG CTGCTCTTTA ACCAGGAGAA TACAACTTAT TCGAAAACGA    1440
TTGCTCTATA TGGATTCTTC TTTAAGATAA AGGGACTCAA ACATGATACT ACCTCTTATA    1500
GTTTTTTACAT GCAGAGAATA AAGCACACAC ACCTGGAATC TCCCTCTGCG GTCTACGAGC    1560
ACAACCACGT CAGCCTGCGA GCGGCACGCC TGGTGAAGTA TGAGATCAGA GCAGAGGCCC    1620
TGGTTGACGG CAAGTGGCAG GAGTTCAGGA CAAACCAGAT CAAGCAGAAG TTTGGGTTGA    1680
CCACGTCATC CTGCAAAAGC CATACCTTGA AAATCCAAAC TGTGGGCATC CCAATCTATG    1740
TAAGTTTTGC ATTCATCTTC CCAGCATCTT GACAGTTTCC AGAAGAATCT ATGGGATTTT    1800
CCCCCCACTG GTCTGCATAA AAGAAAAATAA AATGACATAA AAGGGAGC
```

Seq ID NO: 169 Protein sequence
Protein Accession #: BAB71658.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MIMSNTHKAR LERRVTGSTN RWRLPKQPPS GDLLSLSQMC KALSIDPEEA LRNPDRLCIS    60
QIQKFFPENF KNKDIQSGEA DVILECLGFK WELHQPRLFQ SETLAKLYLK ALAQGTTHPL    120
RELEELLRAQ SPKKTKEKSP AKRIIISLKI NDPLVTKVAF ATALKNLYMS EVEINLEDLL    180
GVLASAHILQ FSGLFQRCVD VMIARLKPST IKKFYEASCK YKEEQLTTGC EKWLEMNLVP    240
LGGTQIHLHK IPQDLLHKVL KSPRLPTFSR FHLLKTMLLW VFLQLNYKIQ AIPTYETVMT    300
FFKSFPENCC FLDRDIGRSL RPLFLCLRLH GITKGKDLEV LRHLNFFPES WLDQVTVNHY    360
HALENGGDMV HLKDLNTQAV RFGLLFNQEN TTYSKTIALY GFFFKIKGLK HDTTSYSFYM    420
QRIKHTDLES PSAVYEHHHV SLRAARLVKY EIRAEALVDG KWQEFRTNQI KQKFGLTTSS    480
CKSHTLKIQT VGIPIYVSPA FIPPAS
```

Seq ID NO: 170 DNA sequence
Nucleic Acid Accession #: NM_007000.1
Coding sequence: 1...777

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGCGTCTG CGGCAGCAGC GGAGGCCGAG AAGGGATCTC CAGTTGTGGT GGGCCTGCTA    60
GTTGTGGGCA ATATCATTAT TCTGCTGTCA GGCCTGTCCC TGTTTGCTGA GACCATATGG    120
GTGACAGCCG ACCAGTACCG TGTATACCCA CTGATGGGAG TCTCAGGCAA GGATGACGTC    180
TTCGCTGGTG CCTGGGATTGC CATCTTCTGC GGCTTCTCCT TCTTCATGGG AGCCAGTTTT    240
```

```
GGTGTGGGTG CCGCACTCTG CCGCCGCCGG TCCATGGTCC TCACGTACCT GGTGCTCATG   300
CTCATCGTCT ACATCTTCGA GTGCGCCTCC TGCATCACGT CCTACACCCA CCGTGACTAC   360
ATGGTGTCCA ACCCATCCCT GATCACCAAG CAGATGCTGA CCTTCTACAG CGCGGACACC   420
GACCAGGGCC AGGAGCTGAC CCGCCTCTGG GACCGCGTCA TGATTGAGCA AGAATGCTGT   480
GGCACATCTG GTCCCATGGA CTGGGTGAAC TTCACGTCAG CCTTCCGGGC GGCCACTCCG   540
GAGGTGGTGT TCCCCTGGCC CCCACTGTGC TGTCGGCCGA CGGGAAACTT CATCCCCCTC   600
AACGAGGAGG GCTGCCGGCCT GGGGCACATG GACTACCTGT TCACCAAGGG CTGCTTCGAA   660
CACATCGGCC ACGCCATCGA CAGCTACACG TGGGGTATCT CGTGGTTTGG GTTTGCCATC   720
CTGATGTGGA CGCTCCCGGT CATGCTGATA GCCATGTATT TCTACACCAT GCTCTGA      720
```

Seq ID NO: 171 Protein sequence
Protein Accession #: NP_008931.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MASAAAAEAE KGSPVVVGLL VVGNIIILLS GLSLFAETIW VIADQYRVYP LMGVSGKDDV   60
FAGAWIAIFC GFSFFMVASF GVGAALCRRR SMVLTYLVLM LIVYIFECAS CITSYTHRDY   120
MVSNPSLITK QMLTFYSADT DQGQELTRLW DRVMIEQECC GTSGPMDWVN FTSAFRAATP   180
EVVFPWPPLC CRRTGNFIPL NEEGCRLGHM DYLFTKGCFE HIGHAIDSYT WGISWFGFAI   240
LMWTLPVMLI AMYFYTML                                                 240
```

Seq ID NO: 172 DNA sequence
Nucleic Acid Accession #: NM_006760.1
Coding sequence: 39..593

```
1          11         21         31         41         51
|          |          |          |          |          |
GAAAGCCTGC CAGCACCTAT TCCACCTCCC AGCCCAGCAT GGCACCCCTG CTGCCCATCC   60
GGACCCTTGCC CTTGATCCTG ATTCTGCTGG CTCTGCTGTC CCCAGGGGCT GCAGACTTCA   120
ACATCTCAAG CCTCTCTGGT CTGCTGTCCC CGGCGCTAAC GGAGAGCCTG CTGGTTGCCT   180
TGCCCCCCTG TCACCTCACA GGAGGCAATG CCACACTGAT GGTCCGGAGA GCCAATGACA   240
GCAAAGTGGT GACGTCCAGC TTTGTGGTGC CTCCGTGCCG TGGGCGCAGG GAACTGGTGA   300
GTGTGGTGGA CAGTGGTGCT GGCTTCACAG TCACTCGGCT CAGTGCATAC CAGGTGACAA   360
ACCTCGTGCC AGGAACCAAA TTCTACATTT CCTACCTAGT GAAGAAGGGG ACAGCCACTG   420
AGTCCAGCAG AGAGATCCCA ATGTCCACAC TCCCTCGAAG GAACATGGAA TCCATTGGGC   480
TGGGTATGGC CCGCACAGGG GGCATGGTGG TCATCACGGT GCTGCTCTCT GTCGCCATGT   540
TCCTGCTGGT GCTGGGCTTC ATCATTGCCC TGGCACTGGG CTCCCGCAAG TAAGGAGGTC   600
TGCCCGGAGC AGCAGCTTCT CCAGGAAGCC CAGGGCACCA TCCAGCTCCC CAGCCCACCT   660
GCTCCCAGGC CCCAGGCCTG TGGCTCCCTT GGTGCCCTGG CCTCCTCCTC CTGCCCTCCT   720
CTCCCCTAGA GCCCCTCTCCT CCCTCTGTCC CTCTCCTTGC CCCCAGTGCC TCACCTTCCA   780
ACACTCCATT ATTCCTCTCA CCCCACTCCT GTCAGAGTTG ACTTTCCTCC CATTTTTACCA   840
CTTTAAACAC CCCCATAACA ATTCCCCCAT CCTTCAGTGA ACTAAGTCCC TATAATAAAG   900
GCTGAGGCTG CATCTGCCAA AAAAAAAAAA AA                                 900
```

Seq ID NO: 173 Protein sequence
Protein Accession #: NP_006751.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MAPLLPIRTL PLILILLALL SPGAADFNIS SLSGLLSPAL TESLLVALPP CHLTGGNATL   60
MVRRANDSKV VTSSFVVPPC RGRRELVSVV DSGAGFTVTR LSAYQVTNLV PGTKFYISYL   120
VKKGTATESS REIPMSTLPR RNMESIGLGM ARTGGMVVIT VLLSVAMFLL VLGFIIALAL   180
GSRK                                                               180
```

Seq ID NO: 174 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..2733

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAAAGTTG GAGTGCTGTG GCTCATTTCT TTCTTCACCT TCACTGACGG CCACGGTGGC   60
TTCCTGGGGA AAAATGATGG CATCAAAACA AAAAAGAAC TCATTGTGAA TAAGAAAAAA   120
CATCTAGGCC CAGTCGAAGA ATATCAGCTG CTGCTTCAGG TGACCTATAG AGATTCCAAG   180
GAGAAAAGAG ATTTGAGAAA TTTTCTGAAG CTCTTGAAGC CTCCATTATT ATGGTCACAT   240
GGGCTAATTA GAATTATCAG AGCAAAGGCT ACCACAGACT GCAACAGCCT GAATGGAGTC   300
CTGCAGTGTA CCTGTGAAGA CAGCTACACC TGGTTTCCTC CCTCATGCCT TGATCCCCAG   360
AACTGCTACC TTCACACGGC TGGAGCACTC CCAAGCTGTG AATGTCATCT CAACAACCTC   420
AGCCAGAGTG TCAATTTCTG TGAGAGAACA AAGATTTGGG GCACTTTCAA AATTAATGAA   480
AGGTTTACAA ATGACCTTTT GAATTCATCT TCTGCTATAT ACTCCAAATA TGCAAATGGA   540
ATTGAAATTC AACTTAAAAA AGCATATGAA AGAATTCAAG GTTTTGAGTC GGTTCAGGTC   600
ACCCAATTTC GAAATGGAAG CATCGTTGCT GGGTATGAAG TTGTTGGCTC CAGCAGTGCA   660
TCTGAACTGC TGTCAGCCAT TGAACATGTT GCCGAGAGGG CTAAGACAGC CCTTCACAAG   720
CTGTTTCCAT TAGAAGACGG CTCTTTCAGA GTGTTCGGAA AAGCCCAGTG TAATGACATT   780
GTCTTTGGAT TTGGGTCCAA GGATGATGAA TATACCCTGC CCTGCAGCAG TGGCTACAGG   840
GGAAACATCA CAGCCAAGTG TGAGTCCTCT GGGTGGCAGG TCATCAGGGA GACTTGTGTG   900
CTCTCTCTGC TTGAAGAACT GAACAAGAAT TTCAGTATGA TTGTAGGCAA TGCCACTGAG   960
GCAGCTGTGT CATCCTTCGT GCAAAATCTT TCTGTCATCA TTCGGCAAAA CCCATCAACC   1020
ACAGTGGGGA ATCGGCTTC GGTGGTGTCG ATTCTGAGCA ATATTTCATC TCTGTCACTG   1080
GCCAGCCATT TCAGGGTGTC CAATTCAACA ATGGAGGATG TCATCAGTAT AGCTGACAAT   1140
ATCCTTAATT CAGCCTCAGT AACCAACTGG ACAGTCTTAC TGCGGGAAGA AAAGTATGCC   1200
AGCTCACGGT TACTAGAGAC ATTAGAAAAC ATCAGCACTC TGGTGCCTCC GACAGCTCTT   1260
CCTCTGAATT TTCTCGGAA ATTCATTGAC TGGAAAGGGA TTCCAGTGAA CAAAAGCCAA   1320
CTCAAAAGGG GTTACAGCTA TCAGATTAAA ATGTGTCCCC AAAATACATC TATTCCCATC   1380
AGAGGCCGTG TGTTAATTGG GTCAGACCAA TTCCAGAGAT CCCTTCCAGA AACTATTATC   1440
```

```
AGCATGGCCT  CGTTGACTCT  GGGGAACATT  CTACCCGTTT  CCAAAAATGG  AAATGCTCAG   1500
GTCAATGGAC  CTGTGATATC  CACGGTTATT  CAAAACTATT  CCATAAATGA  AGTTTTTCCTA  1560
TTTTTTTCCA  AGATAGAGTC  AAACCTGAGC  CAGCCTCATT  GTGTGTTTTG  GGATTTCAGT   1620
CATTTGCAGT  GGAACGATGC  AGGCTGCCAC  CTAGTGAATG  AAACTCAAGA  CATCGTGACG   1680
TGCCAATGTA  CTCACTTGAC  CTCCTTCTCC  ATATTGATGT  CACCTTTTGT  CCCCTCTACA   1740
ATCTTCCCCG  TTGTAAAATG  GATCACCTAT  GTGGGACTGG  GTATCTCCAT  TGGAAGTCTC   1800
ATTTTATGCC  TGATCATCGA  GGCTTTGTTT  TGGAAGCAGA  TTAAAAAAAG  CCAAACTCGT   1860
CACACACGTC  GTATTTGCAT  GGTGAACATA  GCCCTGTCCC  TCTTGATTGC  TGATGTCTGG   1920
TTTATTGTTG  GTGCCACAGT  GGACACCACG  GTGAACCCTT  CTGGAGTCTG  CACAGCTGCT   1980
GTGTTCTTTA  CACACTTCTT  CTACCCTCTCT TTGTTCTTCT  GGATGCTCAT  GCTTGGCATC   2040
CTGCTCGGCTT ACCGGATCAT  CCTCGTGTTC  CATCACATGG  CCCAGCATTT  GATGATGGCT   2100
GTTGGATTTT  GCCTGGGTTA  TGGGTGCCCI  CTCATTATAT  CTGTCATTAC  CATTGCTGTC   2160
ACGCAACCTA  GCAATACCTA  CAAAAGGAAA  GATGTGTGTT  GGCTTAACTG  GTCCAATGGA   2220
AGCAAACCAC  TCCTGGCTTT  TGTTGTCCCT  GCACTGGCTA  TTGTGGCTGT  GAACTTCGTT   2280
GTGGTGCTGC  TAGTTCTCAC  AAAGCTCTGG  AGGCCGACTG  TTGGGGAAAG  ACTGAGTCGG   2340
GATGACAAGG  CCACCATCAT  CCGCGTGGGG  AAGAGCCTCC  TCATTCTGAC  CCCTCTGCTA   2400
GGGCTCACCT  GGGGCTTTGG  AATAGGGACA  ATAGTGGACA  GCCAGAATCT  GGCTTGGCAT   2460
GTTATTTTTG  CTTTACTCAA  TGCATTCCAG  GGATTTTTTA  TCTTTATGCTT TGGAATACTC   2520
TTGGACAGTA  AGCTGCGACA  ACTTCTGTTC  AACAAGTTGT  CTGCCTTAAG  TTCTTGGAAG   2580
CAAACAGAAA  AGCAAACTC  ATCAGATTTA  TCTGCCAAAC  CCAAATTCTC  AAAGCCTTTC   2640
AACCCACTGC  AAAACAAAGG  CCATTATGCA  TTTTCTCATA  CTGGAGATTC  CTCCGACAAC   2700
ATCATGCTAA  CTCAGTTTGT  CTCAAATGAA  TAA
```

Seq ID NO: 175 Protein sequence
Protein Accession #: Eos sequence

```
1          11         21         31         41         51
|          |          |          |          |          |
MKVGVLWLIS  FFTFTDGHGG  FLGKNDGIKT  KKELIVNKKK  HLGPVEEYQL  LLQVTYRDSK   60
EKRDLRNFLK  LLKPPLLWSH  GLIRIIRAKA  TTDCNSLNGV  LQCTCEDSYT  WFPPSCLDFQ   120
NCYLETAGAL  PSCECHLNNL  SQSVNFCERT  KIWGTFKINE  RFTNDLLNSS  SAIYSKYANG   180
IEIQLKKAYE  RIQGFESVQV  TQFRNGSIVA  GYEVVGSSSA  SELLSAIEEV  AEKAKTALHK   240
LFPLEDGSFR  VFGKAQCNDI  VFGFGSKDDE  YTLPCSSGYR  GNITAKCESS  GWQVIRETCV   300
LSLLEELNKN  PSMIVGNATE  AAVSSFVQNL  SVIIRQNPST  TVGNLASVVS  ILSNIGSLSL   360
ASHFRVSNST  MEDVISIADN  ILNSASVTNW  TVLLREEKYA  SSRLLETLEN  ISTLVPPTAL   420
PLNFSRKFID  WKGIPVNKSQ  LKRGYSYQIK  MCPQNTSIPI  RGRVLIGSDQ  PQRSLPETII   480
SMABLTLGNI  LPVSKNGNAQ  VNGPVISTVI  QNYSINEVFL  FFSKIESNLS  QPHCVFWDFS   540
HLQWNDAGCH  LVNETQDIVT  CQCTHLTSFS  ILMSPFVPST  IFPVVKWITY  VGLGISIGSL   600
ILCLIIEALF  WKQIKKSQTS  RTRRICMVNI  ALSLLIADVW  FIVGATVDTT  VNPSGVCTAA   660
VFFTHFFYLS  LFFWMLMLGI  LLAYRIILVP  HHMAQHLMMA  VGFCLGYGCP  LIISVITIAV   720
TQPSNTYKRK  DVCWLNWSNG  SKPLLAFVVP  ALAIVAVNFV  VVLLVLTKLW  RPTVGERLSR   780
DDKATIIRVG  KSLLILTPLL  GLTWGFGIGT  IVDSQNLAWH  VIFALLNAPQ  GFFILCFGIL   840
LDSKLRQLLF  NKLSALSSWK  QTEKQNSSDL  SAKPKFSKPF  NPLQNKGHYA  FSHTGDSSDN   900
IMLTQFVSNE
```

Seq ID NO: 176 DNA sequence
Nucleic Acid Accession #: AB035089.1
Coding sequence: 9845..10219

```
1          11         21         31         41         51
|          |          |          |          |          |
GGGCATGCAG  CCATCGGGGA  AAATCCATAG  TGCAGATAAA  GCAAGGAGGA  AGAAGAAGGA   60
CAGTTCTAGT  AAAAGGGAGA  ACATCAATAT  AGGATGTTTC  TTAGCAATAG  AAAAAGAAGG   120
CCAAGAGGAA  TTAGGGAGAG  AGTTATAAGA  GATCAGCAAG  GGGACAGGGT  TAGATTTGGT   180
TTGGTTTGAA  AGCATACAGT  AAATATGATG  TCTGTCCCTG  GCAGTGTTGG  CAGAGTAGGA   240
AGGAGGAAGG  GAGGCAAGAG  ATAATATCAT  TTTCTCTGTG  CTCCAACTGT  ACTTACATAT   300
GAGACTATTT  CCCTCTCTGC  TTTTCAAACC  TTACTGGAGT  TGTTTTCCCT  CATGAAAACC   360
AAGAAAGGAA  AGCTAGTTAG  TCTTGTTCTG  AGGTTGTTCA  ATGTATACAT  ATCTATATCT   420
GTAGACAGAA  TCCTTGGGAA  TACAGTAATT  GACATATAIT  CTGTTATTTG  ATGCTTGAAA   480
AATCTCCTCC  ACTAACCAGT  TTCCCTATAG  ATTGCCACAA  GCACATAATA  AGAAACAATA   540
AATAAAATGT  TCTCTTGACT  TTGTTACTTA  ACAATGCTGA  GAAAACTTTA  CAGCCTTCAT   600
AAGGAAGTGA  GGTCCAGGAA  AATCTAGGAG  ATATTTCTTA  ACCAATCTAT  AAAGGCATTA   660
GTAATGACAG  GATATTTCCT  GAAAGTGTAA  TTTCCCATTG  AGGATTTGTT  TTTAATTTCT   720
GGATTCCTGG  AGCCAATGAA  GTTGGTGTAT  GTTTATGAAA  TATCAAGAGA  CATAAGTTGG   780
CAAGTGTTCA  TATGCAAAAA  CTTCTTGGAA  TTTCTGAGTT  CTCTGTGGCA  ATATATGACA   840
TCAGGATATG  TCCAGTCTCA  CACACCAGGA  TATGTCCTTT  CTAGCCTGTC  TATCACATGC   900
TAGGAGAACT  ATTTAGGAAC  AGAAAAAAAT  GCCTGAAATG  ATTTCTCATT  TGAACTCATC   960
CAAGCTTTCT  CTAAATTTAA  GCAAACTCCT  GGTCATTTTC  AGTTAGTACC  TTTCCTTAAG   1020
TTCAACCTTC  AGGGCAAACC  TCCGTGCCTC  AGACGTTTAG  CCATAGTCTG  AAATTCTCTT   1080
CCATAGGATTG GTCCCCTGTA  ACCCGGTTT   GTCTCAGCTT  GTTATCCTGT  TTTTTTCTTC   1140
CCTCCATTCG  CAGGATGAGC  TTGTTGCTTC  TGTCCTATGA  GACATTAGAT  TCCTTTTCTT   1200
TGGTACCCGA  GTAAATCCAT  CCTACTCCAA  TAGAGGAAGG  TCCATTTTTG  TCTTATAGCG   1260
CTGGATGCAG  ACTCAGCTGA  GAAGACCATT  ATTCATTTTT  GGAATTCTTT  ATCTCAGATA   1320
TTTCCTCTTC  TTTCTTTTTC  TTCTATCTTT  GGATTTTTAG  TCCATCAACG  CCCCATTAGT   1380
CTATTCCCCG  ACTTCAATCA  GGGAACTTAT  ACCTCTTAAA  CTCATTCAGA  GACTCAAAAC   1440
ATATATATTG  ATACAGGAGA  CCTAAGAAGA  GCATGTCTTG  GGGGTTGAGG  AAACAGGCAG   1500
GTGAGAAATT  TCCAGATTGG  AAACACAGCT  TCCTTTCTCC  CATCCAGCCC  CTACTTTCAG   1560
CCTATGTGTT  TCTGGCACCT  TGTTGTAGAT  AAATCTCCCT  TGACTTGTG   ATGTGCTGAG   1620
AAAACAAACT  CACGGCTGGT  GTTAAAAAGG  GCCCATGACA  ATACCAAGTG  TTGGGGAGAA   1680
TGTGAGAAA   TCAGAACTCT  ATTCACGGTC  GGTTGGAATG  CACACTTGTG  CAGAATTCTA   1740
TGGAGAAGAG  TCTGGCATTT  CCTCAAAATG  TTAACCTGGA  TTTACCATAT  GACCCAGCGA   1800
TTTCATTCAT  AGGTTTATAC  TCAAAGAAA   TGAAGAAATA  TGCCATGCAA  AAAAATGTAC   1860
ATGAAAGGTC  ACAACATCAT  TATTCATAAT  AGTAAAAGGA  TGGAACAAC   ACAAATGTCC   1920
ATCAACTTAT  GATTAAAGAA  AATCTGGTCT  ATTCATAGAA  TGGAATATTA  TTCGACCACA   1980
```

```
AAAAGGAATG ATGTACTGAT CCATGCAATG ATGTGGACAA ACCATGAAAA TAACACTAGA 2040
TTAAAGAAGC CAGTCACAAA AGGACTTACT GTATGATTCC ATTTACCTGA AATGTTTGGA 2100
ATAGGCAAAT CCATAGAAAC AGGAGGTAGA TTCCTGGTTT CCAGGGTCTC CAGGAAGGGA 2160
AGAATGAAGT ACAAGATTTC TTTTGGAGGT AGTGAAATTG TTGTGGAATG AGATCATGAT 2220
GATGATAGCA CAACTTTGTG AATATAATAA AATCATTGAA TTGTACAGTT GAATTTATGG 2280
TATATAAATT ATATGTTAAT AAAAAGGGGG TCCACAAAAC AAACAGCCCC CCACTCTGGT 2340
TGTCAGGGAG ATATTGGATT AAATGGCCTT GGACAACAAC CCCTCTCCCT GGCCACAGAC 2400
ATTCTTCAGA TTACAAGATA TTCCAGGGGA AACACTGGAA TGAGTCTGAA GCCAGGTGCT 2460
AAACAGAAGG ACCATTGAGA AATGTTGTGA TCCTGACAGG TCAAGCAATT TATTTTTCGG 2520
CTTCATTTTT AAATGTAAAA TTAGAAAGCT GCCATTTAAA ATGGCCCGTC TGTTTCAATT 2580
GCTCTTCTCA GTGTCAGCCT GTTAACTCAA TGTGTTAGTC TGTTTTCATG CTGCTGATAA 2640
AAACATACCT GAGACTGGCA AGAAAAAGAG GTTTAATTGG GCTTAGAGTT CCACGTGATT 2700
GGGGAGGCCT CAGAATCACA GTAGGAGGCA AAAGTTATTC TTACATGGTG GCTGCAAGAG 2760
AAGATGAGGA AGAAGCAAAA GAAGAAACCC CTGATAAACC CATCGGATCT CCTGAGGCTT 2820
ATTAACTATC ATGAGAATAG CACAAGAAAG ACCGGCCCCC ATGATTCAAT TACCTCTACC 2880
TGGGTCCCTC CAATAACATG TGGAAATTCT GGTAGATACA ATTCAAGTTG AGATTTGGGT 2940
GGGAACACAG CCAAACCATA TCACTCAGCA AGGCAGATAA CTTTCTCACT GAGCCTATGC 3000
AACAGAAAAC CATCTGGGAT GGTTGTAAGG GGCACAGGAA GTGACTGGTA GGATCACTGC 3060
CAAAGCTGAG CACTCAGGAG AAGGCAATAG AATCCTATTC TCCATAGTAT GCTATAAGAT 3120
ACTGAAGTAC ACTTCTTCAC TATCTCTTTG GACTTAGAAT TAGCACTACA TTCCTTGTTA 3180
TACAGAAAAA TTACTAAGGA AATTCATAGG ATGACAAAAA CTTTCAGAAC TGAAAAACAG 3240
GAAATGTAAG CTTTTTAGTT CTTCGGTATT CGAAGTATGC CTAAAAGACA ATGCAAAATC 3300
CAAGAAAAGA ATGGTGGGGT TTTTGTTTGT TTGGTTTTGT TTTTGTTTTA CAGCTGGAGT 3360
AGAATACAAA GGGATGGAGT TGAAACAAAT GAGAGGGAAAT TGGAATTCTA AACTTATTCT 3420
CATTGGCATT AGAAAGGCAC CTACATGTAT TTCACATGAG CCGGTGACTG CTGACTTGCA 3480
TTCTTATTTT TTCCCTATAG ATTAAAAGGT AGGTACAATG GTAGAACTGT AATCCTGTCC 3540
TTTGTCATAA ATTTTCATAT TCATAAAGGT GAGTGTTAGC CCGCTTGTGA AATCTGAAGT 3600
TGAGTAACTT CAAATACTAA CCACAGAGGG AAAGGCAGCA AGAGGGAGAGG CATAAATTTA 3660
GGATCTCACC CTTCATTCCA CAGACACACA CAGCCTCTCT GCCCACCTCT GCTTCCTCTA 3720
GGAACACAGG TAAGAGCTTC AAGCCTCTCC AGCTTAATAA CATGAATTAT TTTTGAGAAT 3780
AATAATGATA CTGTGTTCTA TATCATGCAT CTCCTGCATT CTGTCTGATT ATATTTTACT 3840
TATTCTGCCA GAGCAAAATT AAAATACCTA TTTCATCTGA TTTGTCCTTT ATCTAAATTG 3900
CTTAGTTCCA AGTAAACCAA GGCACTTTTA GGAACACAGA GGGAGAGTGC CTTGCAGCCA 3960
GAGAGTCTTG AAGGAGATGT CAGGGACGCA TCTTAACAGC TGGTTGGATG TGATCCACAG 4020
AGGTCTCCTG TTAGCATTCA TTGTAAAGCC ATCCTACCTA GCTCTAGTGT AACCAGCAAT 4080
GAAAGAAAGA TAAAGAGGGT CGATTACTTA TTTACAATAG TCTTTAAAAA CGTAGTTTTG 4140
TAAGCCTTCT AATTAGGACA TTAATATATT TAATATATGC ACATTGTAGA AAGATTGAAG 4200
CGTTAAAAAT AAGAGAAAAA CTTTAAATGT CAAAATCTCA CAACCCAGAT ATATCATTTC 4260
TTTAAGAAAA TTGTACTACA AAATACCATT CCATTTATTA AAGTCATTCT GACAGGAATC 4320
TGATGCTTTT CCAGGAGTTC CAGATCACAT CGAGTTCACC ATGAATTCAC TCAGTGAAGC 4380
CAACACCAAG TTCATGTTCG ATCTGTTCCA ACAGTTCAGA AAATCAAAAG AGAACAACAT 4440
CTTCTATTCC CCTATCAGCA TCACATCAGC ATTAGGGATG GTCCTCTTAG GAGCCAAAGA 4500
CAACACTGCA CAACAAATTA GCAAGGTAGC TATCAGCATC ATTACGTTGT CCTGTTGCAG 4560
TTTTTCTCTG GTTCCGTCGG CTAGCACGCA GATGGTAATA GATGTGGTGG TCTGATGGGT 4620
AGCACAGGGG GCTGTGCAGG AATTCCCATA ACTGTGAGAC CACTGACTTA AACAGATCTT 4680
TTGAGTAAAG TTTTCTTGTC CCGCTTCATG TCTCTTCCAG GTTCTTCACT TTGATCAAGT 4740
CACAGAGAAC ACCACAGAAA AAGCTGCAAC ATATCATGTG AGTCACAGAG CACTCTGATT 4800
CAGCTTTAGA TCCCTGAACA GGTCATAGTT TAAACCTGAG ACTTCACAAA AACTAAGAAA 4860
AGGCCAGTTT TAGGGAAAAT CTTGGACACA AAGATTGAGA CATACAGAGT GGGGTTGGCAT 4920
TTCATGCGCAC ATAATTATTA TTCCTCATTT CTGCGGTTACT AAAAGACAGT CAGCACTGTA 4980
CCTCAGAGCA TAGGTCTGGA TCAGGATAGG CTGGGTTCAG ACTCCAGCTT TGCTCTTCAC 5040
AAATGATGAA TAAGAGCAGG ACACAACTGC TCGGAGTCCC AGTGACCTCA TCCCAGAAAA 5100
CTAAGGGTAA GAAAAAATCT GACTCAATAC ATGCAAATAC ATGCAAATGT TTACAACAGT 5160
GCCTTGCCCA TAAAAGTCAT AATAAATGTT ATTATTATTA TAAAGTAGCT ATAATTATAC 5220
TAATCATAAT AATGTGAAAA TAATTTAATT TTCATTGAGT CATTAATGAG ATTCAGAGGA 5280
ATAAGCACAA GTCCAAGTAT ATTTTGGAAA ATGATTGCTA TGGAATATAT TGGTTTAGAG 5340
CCTAATAGT GCAAAATGCT TTGCTGGAAG GTAGAAAGTT CTAGATTTAA ACAGGCTTAG 5400
GTTCAAAACT TGGCACTTCT AATTTATGTC TCTATAAACA GGGTTTTTTT CCCCATTCTC 5460
TGAGCTTTCT TGTGTTCATC TGAATTGAAC TAAAGACTTA GAGTTACCCA TGTAAAGTCC 5520
TTAGCCATGG ACCTGGCATA CACTCTTCTT ACGTGCAGAG AATGACCATC ATGAGGAAAG 5580
AGCCACAGAT CAGTCAATGT GTCCTACAAG ATAATAGCAC CAACAGGTAT AACAGGGCTT 5640
CCTGGCATAA TCTATTTAAA ATATCCAACC TTCAACATAC TCGTATCCTT GATGACTGTT 5700
AGAAGTGAAA TATGGTCCTT GCCCATAAGG AGCTGAGAGT TTAACTGGGA AGCTAAACCT 5760
AACCCTTTAA ACCAACAAGG AGAAAAATCA CTGGTAGACA GCGCTGCATC TTTAGTTCAG 5820
AAGAGAAAAG ATTGCAGTAC GTTAGAGCAA GAAGAATTTT CTGGAAGAAG TCAAATATAA 5880
GGTGGATTTT GAAGGGTATT TGAGGTGAAA TACACCAATT ATCAGGGAAT AACATCAAAG 5940
GTCCTCAATG AGACTACCAG CATTTAGGGA CTGATCTAAC AGACTTAGCA TGGGTTTAGT 6000
ATTTACATTG ATACAGCAAT TGAATGATCT CCTTTTTTGA TGTTTGAAGG TTGATAGGTC 6060
AGGAAATGTT CATCACCAGT TTCAAAAGCT TCTGACTGAA TTCAACAAAT CCACTGATGC 6120
ATATGAGCTG AAGATCGCCA ACAAGCTCTT CGGAGAAAAG ACGTATCAAT TTTTACAGGT 6180
AATTTCACCT GGCCTACCCA CATTTCATTT GCATCCTGAT GTCTGTGTCT CTGAGTGGCC 6240
AAATGGAAGA AAGCAAGGCA GATGAGCCTG GCCGACCCAG GTGGAGAGCA TTTACTCAGA 6300
GTGCATTAGC TCCATTTCCA CAACTCTCCC CCACTGGAGT GTCCCAGACC CCAACGATAC 6360
ATCACTGAAG TGTGGATTTA GGGATAATCT TGTGATAAAA GAGGAGGTTG TGTAATAGAG 6420
TGAGTAAGAG TAATAAGTAA TAAGATACCA TCGATAAACT GGCACTGACT CAGTCACATA 6480
CGATACATCT TGGTGGGAAA TGTATGACTA ATGGGATATT ATTGGAATGG GCAGGCTTGG 6540
GTGAGTTCCT GAGAATAGTT GAGGAAGTAC CAGGGAAATAT TGAATGCACA GGATGAAAGA 6600
CAAAAACAAA GATCAGAAAC ATCATGGTTA AAATTACTGG AGAGAAGCT GAGAAGCAAT 6660
GAATCTCCTT CAGGGAAGCC TGCTCTGCAG TTTGCAAACC ACAGCCTCTT CTGCTTCTGC 6720
CTTTTGCCAA GATGATATTG ACCTTCAGTG ACCTCTTTCT TGTGCCAGCC CACATTCCCC 6780
TTTTGCATTG CCTACATGAC ACCTGTATAA AAATATCCAT GGACAGGAGA TACTGCATCT 6840
ATTCAGGGTC TGGATTCAGC TTACTGTTGT TACAAATAAG TAAGTTTGGT AATATATAGT 6900
TACATAAATT ACTCCTAATT CCTACTTCTT CCTTCATATC TCAAAGGAAT ATTTAGATGC 6960
CATCAAGAAA TTTTACCAGA CCAGTGTGGA ATCTACTGAT TTTGCAAATG CTCCAGAAGA 7020
```

```
AAGTCGAAAG AAGATTAACT CCTGGGTGGA AAGTCAAACG AATGGTAGGA GAGCCACCCA 7080
TTATAGAAAC ACCTTTGAGA AACCTATGCC AGTGAGCCTT GTGCTTGACA CTGCATGGGG 7140
GAACAGGTGT GGGGATTGAG ATGGGTTTGC AGGGAGGGCT GAAGAGGGCA CTCCAGATGA 7200
AGGATTTGTC CAAATGAATA TGAAGAGAGC CTAGGGGAGC CAAGGAGGAA ATCACAGGAA 7260
GCCAATTAGA TGGAAACACA TCTGGAGAAT TATTTGCTTA TGGCCCTGCA TGACAATAGC 7320
TTTGTGGATC CCCTGTCTCC GCTCAGACCT ATTTTGAGAT CATATCCTTT ACTTTAAATC 7380
AGACTCAAAT TTTTATGATG AATATTTAAT AGAAACATT AGAAAGCGTC TCTCGTCTCC 7440
TTTACTAATT GGGAAACAAG CAGCTCTCTG GTAAATCACC CTTTTGTCTC TGAGCTGGAG 7500
CTGCCTGGAT CACATCTGTA GCCAATGTGT TCTGCAGGGA TTATCACAGC TCTCTTCCCC 7560
ATCAAGGGCA AAGAGCTTGA CAAAGTCTCC ATTCTACAGA CATCTTTCTT ACCTCCCACC 7620
TCTCATTACA GGCCAAACTT ACAGCAACTC AACATGAGAG TGAATAGGGA GATACCCCCG 7680
GAAGTAGTGT CTGACAGCAC AGGACATGCG TTTCATATTA CAGAGCTCAA GTCACTCATC 7740
CTAAAATGCA ATCAGGGCCT CCTTCCTCTG AATGGGGACC CCGTAGTTAA AAAAAAATAA 7800
AAGTAGGAAG AGGAGGGAGG GAGAAAGGAA AGACACATGT TGGAAGAGTA GACAAAATCA 7860
GTTTATCAGT ATTCCAAATC AGATGATTGG AGACATTCAT ACACAGAGAA CGTGAACTCC 7920
TTCTCTATCA CAAGAAGTGA TGTCTCCATC AAGGGTAACT TTATACGACT GGAGCCTTGA 7980
AGAAAGCTGC ATCTGGTGAA CCACTGGTCA GTGAGTCTAA CAATTCAAAG ATCAAAGTCA 8040
GTGAGTCTCA AGCAGGGATT TGGGTCAATA ATTAACGATC AGTCACGAAC ATTTGCAAAG 8100
CATCTTCCAG ACAAGCCATT TGTAGCTTGT GTAAAAGACT CTTTTATTCT TTCCCTTGCA 8160
GAAAAAATTA AAAACCTATT TCCTGATGGG ACTATTGGCA ATGATACGAC ACTGGTTCTT 8220
GTGAACGCAA TCTATTTCAA AGGGCAGTGG GAGAATAAAT TTAAAAAAGA AAACACTAAA 8280
GAGGAAAAAT TTTGGCCAAA CAAGGTATTG TCTATATTTT ATTTATATAG TGTAATATGT 8340
TAATACATGG AATGTTAAAC ATTTCTGATG GAATGTAACA TGATAAGTAA AAAATAAAAA 8400
TTGTTCATGT CTGTTATTTT GTTGTTTTAC TCTTATAACT TTATTTAGTT AGGAATACCT 8460
GAAAAACTAT TGTTCTAAC TCATGGAATT CCTGGGTTAT TTCTTAGAAG AAGAAGGATG 8520
TGTTGCTATC TCAATAATAT TATCTTTTTT GTCTTGTGTT TCACGTGTTA TTTGTTGGAC 8580
ACATTGATTT ATTGCAGAAT ACATACAAAT CTGTACAGAT GATGAGGCAA TACAATTCCT 8640
TTAATTTTGC CTTGCTGGAG GATGTACAGG CCAAGGTCCT GGAAATACCA TACAAAGGCA 8700
AAGATCTAAG CATGGATTGTG CTGCTGCCAA ATGAAATCGA TGGTCTGCAG AAGGTAAGAA 8760
CTTGCATCTA CAACTCTTCC TTCTACTGCC GGACATTTTT CCAAAGATAC CAAGTTTAAA 8820
CAAGGTAAAA GCTTATGACC GAGTTGCCTC AAAATGATGA AAAATTCTAA ATGAGGAATG 8880
ATGACTCACC TTCATATTAC AAATATTTGA GCATAGGGCC TGACACAAAC TGAAAGCTTA 8940
GTTTTTGTTT GTTTGTTTGT TTTTATTATT ATTATTATAA TACTTTAAGC TTTAGGGTAC 9000
ATGTGCACAA TGTGCAGGTT AGTTACATAT GTATACATGT GCCATGCTGG TGTGCTGCAC 9060
CCATTAACTC ATCATTTAGC GTTAGGTATA TCTCCTAATG CTATCCCTCC CCCTCCCCC 9120
CACCCCACAA CAGTCCTCAG AGTGTGATGT TACCTTCCTG TGTCCAAGTG TTCTCATTGT 9180
TCAATTCCCA TCTATGATTT AATTCCATCT ATGGCTTAGT TAATGATTAA TTTATTAGAG 9240
TTACATGCAT TGGATATCAA TTTGATGATA TTATTATGCA GCAATTTAAA CTTGACTGGG 9300
AGAAATATAT ACCAATGTGA GGAAAGTTTA CAAATAGGCC GAGTAGAAAA GGGAATACAA 9360
ATTTAGGAAT TTAGGGAATT ACAATTTAAT AATTGCAATG TGTACTAAAT AATGTATACA 9420
GAAAAATATG ATGAGCCTAT TAAAAATTGA CACATGTAGT AGGCTGTTGG CACAAGAAAT 9480
AGTGATACAT ACAGTTCATT GTGTACAAAA TAATGTAATC ATATTTTACA TGTGTATCAT 9540
ACAGTTGTAT ACATACATAT GTACACATAT ACATATACGT AAAAACATGA TTCTGTTTTT 9600
ACATACATGT ATATACACAT ACACATATAA CCCAATGTAT TTATATATTC AGGACTCATA 9660
TTTTACCTAT TAGAATAATA ATGTCTATTA AAGTGAACCT TCTGTATTTC ACATTTATTG 9720
CCAAAATAAC GAATCTCCAC ATAGTCAATT CATTGTTAAG GTGTATTAGA GATCGACAGT 9780
TAGTCATATC AGTTTCTTTT TTCCATTTGT ATAGCTTGAA GAGAAACTCA CTGCTGAGAA 9840
ATTGATGGAA TGGACAAGTT TGCAGAATAT GAGAGAGACA TGTGTCGATT TACACTTACC 9900
TCGGTTCAAA ATGGAAGAGA GCTATGACCT CAAGGACACG TTGAGAACCA TGGGAATGGT 9960
GAATATCTTC AATGGGGATG CAGACCTCTC AGGCATGACC TGGAGCCACG GTCTCTCAGT 10020
ATCTAAAGTC CTACACAAGG CCTTTGTGGA GGTCACTGAG GAGGGAGTGG AAGCTGCAGC 10080
TGCCACGGCT GTAGTAGTAG TCGAATTATC ATCTCCTTCA ACTAATGAAG AGTTCTGTTG 10140
TAATCACCCT TTCCTATTCT TCATAAGGCA AAATAAGACC AACAGCATCC TCTTCTATGG 10200
CAGATTCTCA TCCCCATAGA TGCAATTAGT CTGTCACTCC ATTTAGAAAA TGTTCACCTA 10260
GAGGTGTTCT GGTAAACTGA TTGCTGGCAA CAACAGATTC TCTTGGCTCA TATTTCTTTT 10320
CTATCTCATC TTGATGATGA TAGTCATCAT CAAGAATTTA ATGATTAAAA TAGCCATGCCT 10380
TTCTCTCTTT CTCTTAATAA GCCCACATAT AAATGTACTT TTCCTTCCAG AAAAATTTCC 10440
CTTGAGGAAA AATGTCCAAG ATAAGATGAA TCATTTAATA CCGTGTCTTC TAAATTGAA 10500
ATATAATTCT GTTTCTGACC TGTTTTAAAT GAACCAACCA AAATCATACT TTCTCTTCAA 10560
ATTTAGCAAC CTAGAAACAC ACATTTCTTT GAATTTAGGT GATACCTAAA TCCTTCTTAT 10620
GTTTCTAAAT TTTGTGATTC TATAAACAC ATCATCAATA AAATAATGAC ATAAAATCAT 10680
TTTTGCTTTA CCTGTTTTCT CTCTGGAAAG GGCAAGTGTC CAGTTACACA TAGGAAAGAT 10740
AATTTAGAGA TATATTAATC ATATATAAAG GAAAATTAAA AACAGAGTAG TTCATGATGA 10800
GCCTGGAGTA GAAGGCATAT CCCAGAACAG GAGGAGCCTT GTAAACCACA TAGGAACTTC 10860
CTATTTATG CTAAAGGGAT AAGAAACTCA TTACAGGCTT TGATGGTTGT TTGTCAAAGA 10920
GGGGCATAAA ATTATCATAT CCCACATCTAG AAAATACATC TCTGGCTACG CTGATATCAA 10980
TGGATGCGAG GAAAGAACAG TGTGGTTACC ATATATAAAT TAGGAAATCA TTAGAGTATT 11040
GGGAGTGGGA ATGGAGAGAA AGAAAGAGCC TGGGGGAATT ATTTAGGAAA TAATAGTTAC 11100
AGAAAGACAT CTAAGTTGCT GACCTATCTG ACTGGATGGA TGGAAGAATA TCTTGTTTCT 11160
GAGAGAAAAA AAGACTTTGG GTTTAAATTT GTACTTGATG AATTAAGGTA CTTTTAAATAT 11220
TCAAATGGAT TTGCCTGGCA GGCACTTGAA GATATTAGTC TAAATCTCAG AAACAGAATA 11280
TGATCTGAAG CTCTAAATTT GTGATATTCA ATATAAATAC TTTAGAGTCA TTGGGATAAA 11340
TATGGTAGTT GTAGCTAAAA GCAAAAATAA GATACTAGGG AGAAAGGATA AAGTTAGAAG 11400
AAAGAAGAAT CTAGAATTGA CCTTGAAGTA TATCAGCATG TGTAAAGATC AGGAATTGAT 11460
CATTTTTATT TTCCAGAAAG TAGCTTTTCT TAGGGTTCCA TATTTACTCC CATAGATTCT 11520
TCCC
```

Seq ID NO: 177 Protein sequence
Protein Accession #: BAB21525.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MNSLSEANTK FMFDLFQQFR KSKENNIFYS PISITSALGM VLLGAKDNTA QQISKVLHFD  60
QVTENTTEKA ATYHVDRSGN VHRQFQKLLT EPNKSTDAYE LKIANKLFGE KTYQFLQEYL  120
```

```
DAIKKFYQTS  VESTDFANAP  EESRKKINSW  VESQTNEKIK  NLFPDGTIGN  DTTLVLVNAI   180
YFKGQWENKF  KKENTKEEKF  WPNKNTYKSV  QMMRQYNSFN  FALLEDVQAK  VLEIPYKGKD   240
LSMIVLLPNE  IDGLQKLEEK  LTAEKLMEWT  SLQNMRETCV  DLHLPRFKME  ESYDLKDTLR   300
TMGMVNIFNG  DADLSGMTWS  HGLSVSKVLH  KAFVEVTEEG  VEAAAATAVV  VVELSSPSTN   360
EEPCCNHPFL  FFIRQNKTNS  ILFYGRFSSP
```

```
 1          11         21         31         41         51
 |          |          |          |          |          |
GGAGAGAAGA  AAGGAGGGGG  CAAGGGAGAA  GCTGCTGGTC  GGACTCACAA  TGAAAACGCT    60
CCTTCTTTTG  CTGCTGGTGC  TCCTGGAGCT  GGGAGAGGCC  CAAGGGATCC  TTCACAGGGT   120
GCCCCTCAGG  AGGCATCCGT  CCCCTCAAGAA GAAGCTGCGG  GCACGGAGCC  AGCTCTCTGA   180
GTTCTGGAAA  TCCCATAATT  TGGACATGAT  CCAGTTCACC  GAGTCCTGCT  CAATGGACCA   240
GAGTGCCAAG  GAACCCCTCA  TCAACTACTT  GGATATGGAA  TACTTCGGCA  CTATCTCCAT   300
TGGCTCCCCA  CCACAGAACT  TCACTGTCAT  CTTCGACACT  GGCTCCTCCA  ACCTCTGGGT   360
CCCCTCTGTG  TACTGCACTA  GCCCAGCCTG  CAAGACGCAC  AGCAGGTTCC  AGCCTTCCCA   420
GTCCAGCACA  TACAGCCAGC  CAGGTCAATC  TTTCTCCATT  CAGTATGGAA  CCGGGAGCTT   480
GTCCGGGATC  ATTGGAGCCG  ACCAAGTCTC  TGTGGAAGGA  CTAACCGTGG  TTGGCCAGCA   540
GTTTGGAGAA  AGTGTCACAG  AGCCAGGCCA  GACCTTTGTG  GATGCAGAGT  TTGATGGAAT   600
TCTGGGCCTG  GGATACCCCT  CCTTGGCTGT  GGGAGGAGTG  ACTCCAGTAT  TTGACAACAT   660
GATGGCTCAG  AACCTGGTGG  ACTTGCCGAT  GTTTTCTGTC  TACATGAGCA  GTAACCCAGA   720
AGGTGGTGCG  GGGAGCGAGC  TGATTTTTTG  AGGCTACGAC  CACTCCCATT  TCTCTGGGAG   780
CCTGAATTGG  GTCCCAGTCA  CCAAGCAAGC  TTACTGGCAG  ATTGCACTGG  ATAACATCCA   840
GGTGGGAGGC  ACTGTTATGT  TCTGCTCCGA  GGGGCTGCCAG GCCATTGTGG  ACACAGGGAC   900
TTCCCTCATC  ACTGGCCCTT  CCGACAAGAT  TAAGCAGCTG  CAAAACGCCA  TTGGGGCAGC   960
CCCCGTGGAT  GGAGAATATG  CTGTGGAGTG  TGCCAACCTT  AACGTCATGC  CGGATGTCAC  1020
CTTCACCATT  AACGGAGTCC  CCTATACCCT  CAGCCCAACT  GCCTACACCG  TACTGGACTT  1080
CGTGGATGGA  ATGCAGTTCT  GCAGCAGTGG  CTTTCAAGGA  CTTGACATCC  ACCCTCCAGC  1140
TGGGCCCCTC  TGGATCCTGG  GGGATGTCTT  CATTCGACAG  TTTTACTCAG  TCTTTGACCG  1200
TGGGAATAAC  CGTGTGGGAC  TGGCCCCAGC  AGTCCCCTAA  GGAGGGGCGT  TGTGTCTGTG  1260
CCTGCCTGTC  TGACAGACCT  TGAATATGTT  AGGCTGGGGC  ATTCTTTACA  CCTACAAAAA  1320
GTTATTTTCC  AGAGAATGTA  GCTGTTTCCA  GGGTTGCAAC  TTGAATTAAG  ACCAAACAGA  1380
ACATGAGAAT  ACACACACAC  ACACACATAT  ACACACACAC  CATACACACC  CATACACACA  1440
ACTCCCACCA  CCGTCATGAT  GGAGGAATTA  CGTTATACAT  TCATATTTTG  TATTGATTTT  1500
TGATTATGAA  AATCAAAAAT  TTTCACATTT  GATTATGAAA  ATCTCCAAAC  ATATGCACAA  1560
GCAGAGATCA  TGGTATAATA  AATCCCTTTG  CAACTCCACT  CAGCCCCTGAC  AACCCATCCA  1620
CACACGGCCA  GGCCTGTTTA  TCTACACTGC  TGCCCACTCC  TCTCTCCAGC  TCCACATGCT  1680
GTACCTGGAT  CATTCTGAAG  CAAATTCCGA  GCATTACATC  ATTTTGTCCA  TAAATATTTC  1740
TAACATCCTT  AAATATACAA  TCGGAATTCA  AGCATCTCCC  ATTGTCCCAC  AAATGTTTGG  1800
CTGTTTTTGT  AGTTGGATTG  TTTGTATTAG  GATTCAAGCA  AGGCCCATAT  ATTGCATTTA  1860
TTTGAAATGT  CTGTAAGTCT  CTTTCCATCT  ACAGAGTTTA  GCACATTTGA  ACGTTGCTGG  1920
TTGAAATCCC  GAGGTGTCAT  TTGACATGGT  TCTCTGAACT  TATCTTTCCT  ATAAAATGGT  1980
AGTTAGATCT  GGAGGTCTGA  TTTTTGTGGCA AAAATACTTC  CTAGGTGGTG  CTGGGGTACTT 2040
CTTGTTGCAT  CCTGTCAGGA  GGCAGATAAT  GCTGGTGCCT  CTCTATTGGT  AATGTTAAGA  2100
CTGCTGGGTG  GGTTTGGAGT  TCTTGGCTTT  AATCATTCAT  TACAAAGTTC  AGCATTTT
```

```
 1          11         21         31         41         51
 |          |          |          |          |          |
MKTLLLLLLV  LLELGEAQGS  LHRVPLRRHP  SLKKKLRARS  QLSEFWKSHN  LDMIQFTESC    60
SMDQSAKEPL  INYLDMEYFG  TISIGSPPQN  FTVIFDTGSS  NLWVPSVYCT  SPACKTHSRF   120
QPSQSSTYSQ  PGQSFSIQYG  TGSLSGIIGA  DQVSVEGLTV  VGQQFGESVT  EPGQTFVDAE   180
FDGILGLGYP  SLAVGGVTPV  FDNMMAQNLV  DLPMFSVYMS  SNPEGGAGSE  LIFGGYDHSH   240
FSGSLNWVPV  TKQAYWQIAL  DNIQVGGTVM  FCSEGCQAIV  DTGTSLITGP  SDKIKQLQNA   300
IGAAPVDGEY  AVECANLNVM  PDVTFTINGV  PYTLSPTAYT  LLDFVDGMQF  CSSGFQGLDI   360
HPPAGPLWIL  GDVFIRQFYS  VFDRGNNRVG  LAPAVP
```

```
 1          11         21         31         41         51
 |          |          |          |          |          |
TACGCGCTGC  GGGACCGGCA  GGGGAACGCC  ATCGGGGTCA  CAGCCTGCGA  CATCGACGGG    60
GACGGCCGGG  AGGAGATCTA  CTTCCTCAAC  ACCAATAATG  CCTTCTCGGG  GGTGGCCACG   120
TACACCGACA  AGTTGTTCAA  GTTCCGCAAT  AACCGGTGGG  AAGACATCCT  GAGCGATGAG   180
GTCAACGTGG  CCCGTGGTGT  GGCCAGCCTC  TTTGCCGGAC  GCTCTGTGGC  CTGTGTGGAC   240
AGAAAGGGCT  CTGGACGCTA  CTCTATCTAC  ATTGCCAATT  ACGCCTACGG  TAATGTGGGC   300
CCTGATGCCC  TCATTGAAAT  GGACCCTGAG  GCCAGTGACC  TCTCCCGGGG  CATTCTGCCG   360
CTCAGAGATG  TGGCTGCTGA  GGCTGGGGTC  AGCAAATATA  CAGGGGGCCG  AGGCGGTCAGC  420
GTGGGCCCCA  TCCTCAGCAG  CAGTGCCTCG  GATATCTTCT  GCACAATGA   GAATGGGCCT   480
AACTTCCTTT  TCCACAACCG  GGGCGATGGC  ACCTTTGTGG  ACGCTGCGGC  CAGTGGTGGT   540
GTGGACGACC  CCCACCAGCA  TGGGCGAGGT  GTCGCCCTGG  CTGACTTCAA  CCGTGATGGC   600
AAAGTGGACA  TCGTCTATGG  CAACTGGAAT  GGCCCCACC   GCCTCTATCT  GCAAATGAGC   660
ACCCATGGGA  AGGTCCGCTT  CCGGGACATC  GCCTCACCCA  AGTTCTCCAT  GCCCTCCCCT   720
GTCCGCACGG  TCATCACCGC  CGACTTTGAC  AATGACCAGG  AGCTGGAGAT  CTTCTTCAAC   780
AACATTGCCT  ACGGCAGCTC  CTCAGCCAAC  CGCCTCTTCC  GCGTCATCCG  TAGAGAGCAC   840
GGAGACCCCC  TCATCGAGGA  GCTCAATCCC  GGCACGGCCT  TGGAGCCTGA  GGGCCGGGGC   900
ACAGGGGGTG  TGGTGACCGA  CTTCGACGGA  GACGGGATGC  TGGACCTCAT  CTTGTCCCAT   960
```

```
GGAGAGTCCA TGGCTCAGCC GCTGTCCGTC TTCCGGGGCA ATCAGGGCTT CAACAACAAC   1020
TGGCTGCGAG TGGTGCCACG CACCCGGGTT GGGGCCTTTG CCAGGGAGC TAAGGTCGTG   1080
CTCTACACCA AGAAGAGTGG GGCCCACCTG AGGATCATCG ACGGGGGCTC AGGCTACCTG   1140
TGTGAGATGG AGCCCGTGGC ACACTTTGGC CTGGGGAAGG ATGAAGCCAG CAGTGTGGAG   1200
GTGACGTGGC CAGATGGCAA GATGGTGAGC CGGAACGTGG CCAGCGGGGA GATGAACTCA   1260
GTGCTGGAGA TCCTCTACCC CCGGGATGAG GACACACTTC AGGACCCAGC CCCACTGGAG   1320
ACACCAATGA ATGCATCCAG TTCCCATTCG TGTGCCCTCG AGACAAGCCC GTATGTGTCA   1380
ACACCTATGG AAGCTACAGG TGCCGGACCA ACAAGAAGTG CAGTGGGGC TACGAGCCCA   1440
ACGAGGATGG CACAGCCTGC GTGGGGACTC TCGGCCAGTC ACCGGGCCCC CGCCCCACCA   1500
CCCCCACCGC TGCTGCTGCC ACTGCCGCTG CTGCTGCCGC TGCTGGAGCT GCCACTGCTG   1560
CACCGGTCCT CGTAGATGGA GATCTCAATC TGGGGTCGGT GGTTAAGGAG AGCTGCGAGC   1620
CCAGCTGCTG AGCAGGGGTG GGACATGAAC CAGCGGATGG AGTCCAGCAG GGGAGTGGGA   1680
AAGTGGGCTT GTGCTGCTGC CTAGACAGTA GGGATGTAAA GGCCTGGGAG CTAGACCCTC   1740
CCCAAGCCCA TCCATGCACA TTACTTAGCT AACAATTAGG GAGACTCGTA AGGCCAGGCC   1800
CTGTGCTGGG CACATAGCTG TGATCACAGC AGACAGGGTC GCTGCCCTGA TGGCGCTTAC   1860
ATTCCAGTCG GTCTAATGAC CATATCTTAG GACACAGATG TGCCCAGGGA GGTGGTGTCA   1920
CTGCACAGGA AGTATGAGGA CTTTAGTGTC CTGAGTTCAA ATCCTGATTC AGGAACTCAC   1980
AAAGCTATGT GACCTTACAC CAGTCACTTA ACTTGTTAGC CATCCATTAT CGCATCTGCA   2040
AAATGGGGAT TAAGAATAGA ATCTTGGGGT TAGTGTGGAG ATTAGATTAA ATGTATGTAA   2100
GACACTTGGC ACAAAACCTG GCACATAGTA AAGGCTCAAT AAAAACAAGT GCCTCTCACT   2160
GGGCTTTGTC AACACGTG
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MDPEASDLSR GILALRDVAA EAGVSKYTGG RGVSVGPILS SSASDIFCDN ENGPNPLFHN   60
RDDGTFVDAA ASAGVDDPHQ EGRGVALADF NRDGKVDIVY GNWNGPHRLY LQMSTHGKVR  120
FRDIASPKFS MPSPVRTVIT ADFDNDQELE IFFNNIAYRS SSANRLPKVI RREHGDPLIE  180
ELNPGDALEP EGRGTGGVVT DFDGDGMLDL ILSHGESMAQ PLSVFRGNQG FNNNWLRVVP  240
RTRVGAPARG AKVVLYTKKS GAHLRIIDGG SGYLCEMEPV AHPGLGKDEA SSVEVTWPDG  300
KMVSRNVASG EMNSVLEILY PRDEDTLQDP APLETPMNAS SSHSCALETS PYVSTPMEAT  360
GAGPTRSAVG ATSPTRMAQP AWGLSASHRA PAPPPPPLLL PLPLLLPLLE LPLLHRSS
```

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGTCCAGGA TGTTACCGTT CCTGCTGCTG CTCTGGTTTC TGCCCATCAC TGAGGGGTCC   60
CAGCGGCTG AACCCATGTT CACTGCAGTC ACCAACTCAG TTCTGCCTCC TGACTATGAC  120
AGTAATCCCA CCCAGCTCAA CTATGGTGTG GCAGTTACTG ATGTGGACCA TGATGGGGAC  180
TTTGAGATCG TCGTGGCGGG GTACAATGGA CCCAACCTGG TTCTGAAGTA TGACCGGGCC  240
CAGAAGCGGC TGGTGAACAT CGCGGTCGAT GAGCGGCAGCT CACCCTACTA CGCGCTACGG  300
GACCGGCAGG GGAACGCCAT CGGGGTCACA GCCTGCGACA TCGACGGGGA CGGCGGGGAG  360
GAGATCTACT TCCTCAACAC CAATAATGCC TTCTCGGGGG TGGCCACGTA CACCGACAAG  420
TTGTTCAAGT TCCGCAATAA CCGGTGGGAA GACATCCTGA GCGATGAGGT CAACGTGGCC  480
CGTGGTGTGG CCAGCCTCTT TGCCGGACGC TCTGTGGCCT GTGTGGACAG AAAGGGCTCT  540
GGACGCTACT CTATCTACAT TGCCAATTAC GCCTACGGTA ATGTGGGCCC TGATGCCCTC  600
ATTGAAATGG ACCCTGAGGC CAGTGACCTC TCCCGGGGCA TTCTGCGGCT CAGAGATGTG  660
GCTGCTGAGG CTGGGGTCAG CAAATATACA GGGGGCCGAG GCGTCAGCGT GGGGCCCATC  720
CTCAGCAGCA GTGCCTCGGA TATCTTCTGC GACAATGAGA ATGGGCCTAA CTTCCTTTTC  780
CACAACCGGG GCGATGGCAC CTTTGTGGAC GCTGCGGCCA GTGCTGGTGT GGACGACCCC  840
CACCAGCATG GGCGAGGTGT CGCCCTGGCT GACTTCAACC GTGATGGCAA AGTGGACATC  900
GTCTATGGCA ACTGGAATGG CCCCCACCGC CTCTATCTGC AAATGAGCAC CCATGGGAAG  960
GTCCGCTTCC GGGACATCGC CTCACCCAAG TTCTCCATGC CCTCCCCTGT CCGCACGGTC  1020
ATCACCGCCG ACTTTGACAA TGACCAGGAG CTGGAGATCT TCTTCAACAA CATTGCCTAC  1080
CGCAGCTCCT CAGCCAACCG CCTCTTCCGC GTCATCCGTA GAGAGCACGG AGACCCCCTC  1140
ATCGAGGAGC TCAATCCCGG CGACGCCTTG GAGCCTGAGG GCCGGGGCAC AGGGGGTGTG  1200
GTGACCGACT TCGACGGAGA CGGGATGCTG GACCTCATCT TGTCCCATGG AGAGTCCATG  1260
GCTCAGCCGC TGTCCGTCTT CCGGGGCAAT CAGGGCTTCA ACAACAACTG GCTGCGAGTG  1320
GTGCCACGCA CCCGGGTTGG GGCCTTTGCC AGGGGAGCTA AGGTCGTGCT CTACACCAAG  1380
AAGAGTGGGG CCCACCTGAG GATCATCGAC GGGGGCTCAG GCTACCTGTG TGAGATGGAG  1440
CCCGTGGCAC ACTTTGGCCT GGGGAAGGAT GAAGCCAGCA GTGTGGAGGT GACGTGGCCA  1500
GATGGCAAGA TGGTGAGCCG GAACGTGGCC AGCGGGGAGA TGAACTCAGT GCTGGAGATC  1560
CTCTACCCCC GGGATGAGGA CACACTTCAG GACCCAGCCC CACTGGAGTG TGGCCAAGGA  1620
TTCTCCCAGC AGGAAAATGC CCATTGCATG GACACCAATG AATGCATCCA GTTCCCATTC  1680
GTGTGCCCTC GAGACAAGCC CGTATGTGTC AACACCTATG GAAGCTACAG GTGCCGGACC  1740
AACAAGAAGT GCAGTCGGGG CTACGAGCCC AACGAGGATG GCACAGCCTG CGTGGGGACT  1800
CTCGGCCAGT CACCGGGCCC CCGCCCCACC ACCCCCACCG CTGCTGCTGC CACTGCCGCT  1860
GCTGCTGCCG CTGCTGGAGC TGCCACTGCT GCACCGGTCC TCGTAGATGG AGATCTCAAT  1920
CTGGGGTCGG TGGTTAAGGA GAGCTGCGAG CCCAGCTGCT GAGCAGGGGT GGGACATGAA  1980
CCAGCGGATG GAGTCCAGCA GGGGAGTGGG AAAGTGGGCT TGTGCTGCTG CCTAGACAGT  2040
AGGGATGTAA AGGCCTGGGA GCTAGACCCT CCCCAAGCCC ATCCATGCAC ATTACTTAGC  2100
TAACAATTAG GGAGACTCGT AAGGCCAGGC CCTGTGCTGG GCACATAGCT GTGATCACAG  2160
CAGACAGGGT CGCTGCCCTG ATGGCGCTTA CATTCCAGTC GGTCTAATGA CCATATCTTA  2220
GGACACAGAT GTGCCCAGGG AGGTGGTGTC ACTGCACAGG AAGTATGAGG ACTTTAGTGT  2280
CCTGAGTTCA AATCCTGATT CAGGAACTCA CAAAGCTATG TGACCTTACA CCAGTCACTT  2340
AACTTGTTAG CCATCCATTA TCGCATCTGC AAAATGGGGA TTAAGAATAG AATCTTGGGG  2400
TTAGTGTGGA GATTAGATTA AATGTATGTA AGACACTTGG CACAAAACCT GGCACATAGT  2460
AAAGGCTCAA TAAAAACAAG TGCCTCTCAC TGGGCTTTGT CAACACG
```

Seq ID NO: 183 Protein sequence
Protein Accession #: CAC08451

```
      1          11         21         31         41         51
      |          |          |          |          |          |
    MSRMLPFLLL LWFLPITEGS QRAEPMFTAV TNSVLPPDYD SNPTQLNYGV AVTDVDHDGD    60
    FEIVVAGYNG PNLVLKYDRA QKRLVNIAVD ERSSPYYALR DRQGNAIGVT ACDIDGDGRE   120
    BIYFLNTNNA PSGVATYTDK LFKFRNNRWE DILSDEVNVA RGVASLFAGR SVACVDRKGS   180
    GRYSIYIANY AYGNVGPDAL IEMDPEASDL SRGILALRDV AAEAGVSKYT GGRGVSVGPI   240
    LSSSASDIFC DNENGPNFLF HNRGDGTFVD AAASAGVDDP BQHGRGVALA DFNRDGKVDI   300
    VYGNWNGPHR LYLQMSTHGK VRFRDIASPK FSMPSPVRTV ITADFDNDQE LEIFFNNIAY   360
    RSSSANRLFR VIRREHGDPL IEELNPGDAL EPEGRGTGGV VTDFDGDGML DLILSHGESM   420
    AQPLSVFRGN QGFNNNWLRV VPRTRFGAFA RGAKVVLYTK KSGAHLRIID GGSGYLCEME   480
    PVAHFGLGKD EASSVEVTWP DGKMVSRNVA SGEMNSVLBI LYPRDEDTLQ DPAPLECGQG   540
    FSQQENGHCM DTNECIQFPF VCPRDKPVCV NTYGSYRCRT NKKCSRGYEP NEDGTACVGT   600
    LGQSPGPRPT TPTAAAATAA AAAAAGAATA APVLVDGDLN LGSVVKESCE PSC
```

Seq ID NO: 184 DNA sequence
Nucleic Acid Accession #: FGENESHB
Coding sequence: 1..4794

```
      1          11         21         31         41         51
      |          |          |          |          |          |
    ATGGCGTGTC CGGGAGGACT CCCAGCCCGT TGCTCTGGTT GGATGGGACT GGGTGGGCCC    60
    AGCGGCTCCT CCCCAGCATC CCCTCCCCAT TCCTCCTCCA GGTACAATGG ACCCAACCTG   120
    GTTCTGAAGT ATGACCGGGC CCAGAAGCGG CTGGTGAACA TCGCGGTCGA TGAGCGCAGC   180
    TCACCCTACT ACGCGCTGCG GGACCGGCAG GGGAACGCCA TCGGGGTCAC AGCCTGCGAC   240
    ATCGACGGGG ACGGCCGGGA GGAGATCTAC TTCCTCAACA CCAATAATGC CTTCTCGGGC   300
    CACAGCAGCT CAGCGCAGGT CCCTTCTGGG CTCCACAGAA ACAGGCCTGT GCTGAAGCCT   360
    CCACCTACAA CCCCTGCAGG CCTCCTGGGT CTGCCTCCAC TCAGCGGAAG GGACTTTTCC   420
    TCCTCCCTGG GTCAGGCTTC TCCGGACAGC AGGCAGGGAG AGAGGGTGCC GGTTCCCTGC   480
    TGTCGGGGTG GACTGAGACC TACCCATGAA CCAGAACCAT TTCTTCTGAG ACCCAAATCA   540
    GGGGTGGCCA CGTACACCGA CAAGTTGTTC AAGTTCCGCA ATAACCGGTG GGAAGACATC   600
    CTGAGCGATG AGGTCAACGT GGCCCGTGGT GTGGCCAGCC TCTTTGCCGG ACGCTCTGTG   660
    GCCTGTGTGG ACAGAAAGGG CTCTGGACGC TACTCTATCT ACATTGCCAA TTACGCCTAC   720
    GGTAATGTGG GCCCTGATGC CCTCATTGAA ATGGACCCTG AGGCCAGTGA CCTCTCCCGG   780
    GGCATTCTGG CGCTCAGAGA TGTGGCTGCT GAGGCTGGGG TCAGCAAATA TACAGAAGGC   840
    TTCTCCCACA CTGCCTCTCC AAGCATTGGT GAGATATCTG GCAGAACCGA GGAGCGGGAA   900
    GGAGGAGACC CAGAGGAGGC AGATGAGGAG CACAGTGGGG ATGGAAGCAC CAGCCAACTG   960
    TGCCGGCTGG GCTGGAAGGA CGGGCAGTTC AAGGAAGAAG CAGCAGCTTT GGTGGAGGAA  1020
    CAGAGGGAGG CTGGGGCAGC TGGCGTGCCC AGAGGACGTG TTCGAACAGC TCTGCAGACT  1080
    TCCAAAAGCC ATTTGGCTGA CAAGAACCTA TTTGGCCCAC CATGTTACTA TTCTGTCTGC  1140
    GCGCCTTCTC CAGCCCACCC TTTCCCTGCC CGCCAAGCCC CCAACACTA CCCTGTAGCC  1200
    CCCCTTGTCA CTCAGCTAAT GACACATGGA CGTCTGGCTG GAAAACTAGC CCGGAGTGTC  1260
    CCCCACCCCC GAGCCCCAGG AATGGACCCC AAATGTAAGG GCCGCCATGC TGAGCCCGGC  1320
    CTGATGGCTG AGGCTTTGGG CGCGTGGCCA GCGCTCAGCA CCACTGTGGT GCCAGGGGGC  1380
    CTGAGAAGCT GGGAGGAAAG CAGGCAGAAG GGGCAGGCCA TGTCCAGATG TGCACTCAGG  1440
    GAGCTGGGAG GTCCCTGGAG CCAAGCCACA CAGCACCTGC CTGCTAGAGA GCTGTATGAC  1500
    CTGGGAGAAC CTCCCATTTT ACAAAGAACA GACGGAGATC CAGGGAGGAG AAGGGACTCG  1560
    CCCAAGGTCA CACAGGAGTG CCATCTAGTG GCCACCATGC CAGCTCTCGG GGGACTCGAG  1620
    GGCCCCGGGA GGGTGGCCAA GCGAGAGATT GGGAGAGAGA CTGGGGCAGT AGGAAGACCA  1680
    CTCTCCCATC CCCTGGTCCC CAACTTCCCC AGCTGCTTGA GGCCTCTTGA AGCCGGGACA  1740
    GTGCCGGGAG CTGCCCTGCC TGGGAATCCT GGGAACTGGG TTCTGGACAT GGCCAAGGCC  1800
    CTGGCCGTGGA ACCAGATGGA AAAAGAGGAG GGGAAGATTC ATGGAGACCA TGAGCCCAGA  1860
    TTTAGGCTCA GGAAAGCACG GGAAGCAGAA TTCCCCCAG GCTCCTCTGA GGAGCCTCTG  1920
    CTGCAGTTCC CCTCAGGCCT CAGAGGCAGC CCTGTCCTCC AGGTGGGCCT GGGGCTTGCT  1980
    TCTGCCACTC ACTGTGGGTC GATGTCTTTT CTAGGGGGCC GAGGCGTCAG CGTGGGGCCC  2040
    ATCCTCAGCA GCAGTGCCTC GGATATCTTC TGCGACAATG AGAATGGGCC TAACTTCCTT  2100
    TTCCACAACC GGGGCGATGG CACCTTTGTG GACGCTGCGG CCAGTGCTAA GGTCGTTTA   2160
    GCCTTCATCG TTCACCTCAA ATATCACCTC TGCAGAGATT TTCCTCACTC CCTGTGCCAC  2220
    CTAGCAGAAA CTGGTCCTTC CTCCTCCTGC TGCCCGTGGC ATGCACGTCT TCTTCAGGCT  2280
    CCACATTGCC ATCATGGTTT GTCTATGAGC TTTACAAGGA CGGGTCACG GTTCTATTCA  2340
    TTCTTGACGC AAGGCTTGGC CTCCAGTGCC CACCGGAGGA CACTCAGCCT CCAGGGTTCT  2400
    CAGGGGGCCC CACCCTGCCT TCTGGCAAGA GCTCCCTGTG TCCTGGGGTC TCTGATCCCC  2460
    ACTGCCTATT ACATTGTCCT GTGGTCTGCC ATCCCAGAGA GCCTGATGAC CCACAGCTAT  2520
    TTGTCCTCTG AAAGAGTCAA CGTGGGTGTG GACGACCCCC ACCAGCATGG GCGAGGTGTC  2580
    GCCCTGGCTG ACTTCAACCG TGATGGCAAA GTGGACATCG TCTATGGCAA CTGGAATGGC  2640
    CCCCACCGCC TCTATCTGCA AATGAGCACC CATGGGAAGG TCCGCTTCCG GGACATCGCC  2700
    TCACCCAAGT TCTCCATGCC CTCCCCTGTC CGCACGGTCA TCACCGCCGA CTTTGACAAT  2760
    GACCAGGAGC TGGAGATCTT CTTCAACAAC ATTGCCTACC GCAGCTCCTC AGCCAACCGC  2820
    CTCTTCCGAT GCTCCATCCT GGCTCGTGGC TCTTCATCCT TGACAGCTGG TGGGAGGAAC  2880
    GGTCAGGGAG AAGGTTTAAG AATCAGAAGG GGAGGGTTCC CAGGGCCAGG GGGTCAGGCC  2940
    AAGGTCAACA CAGGTCCCCT GATGAAGAAA CAGAAAGGAA GGAAGGACGA GGACTGGGCA  3000
    AGAGGCTGTG GGAATGCAGG GCAAAGCCTG GCCAAGGAGC CGGCCTCTGC TATTGCAGGG  3060
    AAAGGGAAGG GAAATGTGGC CCAAAGTGTG CCCAGAAACG AAGCGCCACA AGATACAAAG  3120
    CCACACTACC ACAAAAAGGG GCTACAGGGT CCAATCACTA CCAGGAAAAG GGGCTACGGG  3180
    GTCCAATCAC TACCAGGAAA AGGGGCTACG GGGTCCAATC ACTACCAGGA AAAGGGGCTA  3240
    CGGGGTCCAA TCACTACCAG GAAAAGGGGC TACGGGGTCC AATCACTACC AGGAAAAGGG  3300
    GCTACGGGCT CCAATCACTA CCAGGAAAAG GGGCTACAGG GTCCAATCAC TACCAGGAAA  3360
    AGGGGCTACG GGCTCCAATC ACTACCAGGA AAAGGGGCTA CAGGGTCCAA TCACTACCAC  3420
    AGAAAGGGGC TACGGGCTCC AATCACTACC AGGAAAAGGG GCTACGGGGT CCAATCACTA  3480
    CCAGGAAAAG GGGCTACAGG GTCCAATCAC TACCAGGAAA AGGGGCTACG GGTCCAATC   3540
    ACTACCAGGA AAAGGGGCTA CGGGCTCCAA TCACTACCAG GAAAAGGGGC TACGGGGTCC  3600
    AATCACTACC AGGAAAAGGG GCTACAGGGT CCAATCACTA CCAGGAAAAG GGGCTACAGG  3660
```

1027

```
GTCCAATCAC TACCACAGAA AGGGGCTACG GGCTCCAATC ACTACCAGGA AAAGGGGCTA   3720
CGGGGTCCAA TCACTACCAG GAAAAGGGGC TACGGGCTCC AATCACTACC AGGAAAAGAG   3780
GCTATGGGGT CCAATCACTA CCAGGAAAAG GGGCTACGGG CTCCAATCAC TACCAGGAAA   3840
AGGGGCTATG GGGTCCAATC ACTACCACAG AAAGGGGCTA CGGGGTCCAA CGTCATCCGT   3900
AGAGAGCACG GAGACCCCCT CATCGAGGAG CTCAATCCCG GCGACGGCTT GGAGCCTGAG   3960
GGCCGGGGCA CAGGGGGTGT GGTGACCGAC TTCGACGGAG ACGGGGATGCT GGACCTCATC   4020
TTGTCCCATG GAGAGTCCAT GGCTCAGCCG CTGTCCGTCT TCCGGGGCAA TCAGGGCTTC   4080
AACAACAACT GGCTGCGAGT GGTGCCACGC ACCCGGTTTG GGGCCTTTGC CAGGGGAGCT   4140
AAGGTCGTGC TCTACACCAA GAAGAGTGGG GCCCACCTGA GGATCATCGA CGGGGGCTCA   4200
GGCTACCTGT GTGAGATGGA GCCCGTGGCA CACTTTGGCC TGGGGAAGGA TGAAGCCAGC   4260
AGTGTGGAGG TGACGTGGCC AGATGGCAAG ATGGTGAGCC GGAACGTGGC CAGCGGGGAG   4320
ATGAACTCAG TGCTGGAGAT CCTCTACCCC CGGGGATGAGG ACACACTTCA GGACCCAGCC   4380
CCACTGGAGT GTGGCCAAGG ATTCTCCCAG CAGGAAAATG GCCATTGCAT GGACACCAAT   4440
GAATGCATCC AGTTCCCATT CGTGTGCCCT CGAGACACAG CCGTATGTGT CAACACCTAT   4500
GGAAGCTACA GGTGCCGGAC CAACAAGAAG TGCAGTCGGG GCTACGAGCC CAACGAGGAT   4560
GGCACAGCCT GCGTGGGTAC TGAGCTAGGC TCTAGGCATA CAATGACGTG GAAACCAAGG   4620
CCCAAAAAGG AGCTGCAACT TTCCCAAGGC ATCTGCACCC CCGTCTGGTC CTTTTTCCTG   4680
CCGGGGTTGCC GGCTGCTCCT CAAAAGAGCT CAGCTCCAGG CTGCTCCCAG CACCCCTTCTC   4740
CAGAAAGCTC CAGGTATTCC AGAAGCCCAA GTGTATGAAC AAGATCAGGA ATAA
```

Seq ID NO: 185 Protein sequence
Protein Accession #: FGENESHH

```
1          11         21         31         41         51
|          |          |          |          |          |
MACPGGLPAR CSGWMGLGGP SGSSPASPPH SSSHYNGPNL VLKYDRAQKR LVNIAVDERS     60
SPYYALRDRQ GNAIGVTACD IDGDGREEIY FLNTNNAPSG HSSSAQVPSG LHRNRPVLKP    120
PPTTPAGLLG LPPLSGRDFS SSLGQASFDS RQGERVPVPC CRGGLRPTHE PEPFLLRPKS    180
GVATYTDKLF KFRNNRWEDI LSDEVNVARG VASLFAGRSV ACVDRKGSGR YSIYIANYAY    240
GNVGPDALIE MDPEASDLSR GILALRDVAA EAGVSKYTEG FSHTASPSIG EISGRTEERE    300
GGDPEEADEE HSGDGSTSQL CRLGWKDGQF KEEAAALVEE QREAGAAGVP RGRVRTALQT    360
SKSHLADKNL FGPPCYYSVC APSPAHPFPA RQAPQHYPVA PLVTQLMTHG RLAGKLARSV    420
PHPRAPGMDP KCKGRHAEPG LMAEALGAWP ALSTTVVPGG LRSWEESRQK GQAMSRCALR    480
ELGGPWSQAT QHLPARELYD LGEPPILQRT DGDPGRRRDS PKVTQECHLV ATMPALGGLE    540
GPGRVAKREI GRETGAVGRP LSHDLVPNFP SCLRPLEAGT VPGAALPGNP GNWVLDMAKA    600
LAWNQMEKEE GKIHGDHEPR PRLRKAREAE FPPGSSEEPL LQPPSGLRGS PVLQVGLGLA    660
SATHCGSMSF LGGRGVSVGP ILSSSASDIF CDNENGPNFL FHNRGDGTFV DAAASAERRL    720
AFIVHLKYHL CRDFPHSLCH LAETGPSSSC CPWHARLLQA PHCHHGLSMS FTRTGSRFYS    780
FLTQGLASSA HRRTLSLQGS QGAPPCLLAR APCVLGSLIP TAYYIVLWSA IPESLMTHSY    840
LSSERVNVGV DDPHQHGRGV ALADFNRDGK VDIVYGNWNG PHRLYLQMST HGKVRFRDIA    900
SPKFSMPSPV RTVITADFDN DQELEIPFNN IAYRSSSANR LFRCSILARG SSSLTAGGEN    960
GQGEGLRIRR GGFPGPGGQA KVNTGPLMKK QKGRKDEDWA RGCGNAGQSL AKEPASAIAG   1020
KGKGNVAQSV PRTQAPQDTK PHYHKKGLQG PITTRKRGYG VQSLPGKGAT GSNHYQEKGL   1080
RGPITTRKRG YGVQSLPGKG ATGSNHYQEK GLQGPITTRK RGYGLQSLPG KGATGSNHYH   1140
RKGLRAPITT RKRGYGVQSL PGKGATGSNH YQEKGLRGPI TTRKRGYGLQ SLPGKGATGS   1200
NHYQEKGLQG PITTRKRGYR VQSLPQKGAT GSNHYQEKGL RGPITTRKRG YGLQSLPGKE   1260
AMGSNHYQEK GLRAPITTRK RGYGVQSLPQ KGATGSNVIR REHGDPLIEE LNPGDALEPE   1320
GRGTGGVVTD FDGDGMLDLI LSRHGESMAQP LSVFRGNQGF NNNWLRVVPR TRFGAFARGA   1380
KVVLYTKKSG AHLRIIDGGS GYLCSMEPVA HFGLGKDEAS SVEVTWPDGK MVSRNVASGE   1440
MNSVLEILYP RDEDTLQDPA PLECGQGFSQ QENGHCMDTN ECIQPPFVCP RDKPVCVNTY   1500
GSYRCRTNKK CSRGYEPNED GTACVGTELG SRHTMTWKPR PKKELQLSQG ICTPVWSFFL   1560
PGCRLLLKRA QLQAAPSTLL QKAPGIPEAQ VYEQDQE
```

Seq ID NO: 186 DNA sequence
Nucleic Acid Accession #: NM_000584.1
Coding sequence: 75..374

```
1          11         21         31         41         51
|          |          |          |          |          |
AGCAGAGCAC ACAAGCTTCT AGGACAAGAG CCAGGAAGAA ACCACCGGAA GGAACCATCT     60
CACTGTGTGT AAACATGACT TCCAAGCTGG CCGTGGCTCT CTTGGCAGCC TTCCTGATTT    120
CTGCAGCTGT GTGTGAAGGT GCAGTTTTGC CAAGGAGTGC TAAAGAACTT AGATGTCAGT    180
GCATAAAGAC ATACTCCAAA CCTTTCCACC CCAAATTTAT CAAAGAACTG AGAGTGATTG    240
AGAGTGGACC ACACTGCGCC AACACAGAAA TTATTGTAAA GCTTTCTGAT GGAAGAGAGC    300
TCTGTCTGGA CCCCAAGGAA AACTGGGTGC AGAGGGTTGT GGAGAAGTTT TTGAAGAGGG    360
CTGAGAATTC ATAAAAAAAT TCATTCTCTG TGGTATCCAA GAATCAGTGA AGATGCCAGT    420
GAAACTTCAA GCAAATCTAC TTCAACACTT CATGTATTGT GTGGGTCTGT TGTAGGGTTG    480
CCAGATGCAA TACAAGATTC CTGGTTAAAT TTGAATTTCA GTAAACAATG AATAGTTTTT    540
CATTGTACCA TGAAATATCC AGAACATACT TATATGTAAA GTATTATTTA TTTGAATCTA    600
CAAAAAACAA CAAATAATTT TTAAATATAA GGATTTTCCT AGATAGTGCA CGGGGAGAATA    660
TACAAATAGC AAAATTGAGC CAAGGGCCAA GAGAAATATCC GAACTTTAAT TTCAGGAATT    720
GAATGGGTTT GCTAGAATGT GATATTTGAA GCAATCACATA AAAATGATGG GACAATAAAT    780
TTTGCCATAA AGTCAAATTT AGCTGGAAAT CCTGGATTTT TTTCTGTTAA ATCTGGCAAC    840
CCTAGTCTGC TAGCCAGGAT CCACAAGTCC TTGTTCCACT GTGCCTTGGT TTCTCCTTTA    900
TTTCTAAGTG GAAAAGTAAT TAGCCACCAT CTTACCTCAC AGTGATGTTG TGAGGACATG    960
TGGAAGCACT TTAAGTTTTT TCATCATAAC ATAAATTATT TTCAAGTGTA ACTTATTAAC   1020
CTATTTATTA TTTATGTATT TATTTAAGCA TCAAATATTT GTGCAAGAAT TTGGAAAAAT   1080
AGAAGATGAA TCATTGATTG AATAGTTATA AAGATGTTAT AGTAAATTTA TTTTATTTTA   1140
GATATTAAAT GATGTTTTAT TAGATAAATT TCAATCAGGG TTTTTAGATT AAACAAAGAA   1200
ACAATTGGGT ACCCAGTTAA ATTTTCATTT CAGATAAACA ACAATAATT TTTTAGTATA   1260
AGTACATTAT TGTTTATCTG AAAGTTTTAA TTGAACTAAC AATCCTAGTT TGATACTCCC   1320
AGTCTTGTCA TTGCCAGCTG TGTTGGTAGT GCTGTGTTGA ATTACGGAAT AATGAGTTAG   1380
AACTATTAAA ACAGCCAAAA CTCCACAGTC AATATTAGTA ATTTCTTGCT GGTTGAAACT   1440
TGTTTATTAT GTACAAATAG ATTCTTATAA TATTATTTAA ATGACTGCAT TTTTAAATAC   1500
```

```
      AAGGCTTTAT ATTTTTAACT TTAAGATGTT TTTATGTGCT CTCCAAATTT TTTTTACTGT  1560
      TTCTGATTGT ATGGAAATAT AAAAGTAAAT ATGAAACATT TAAAATATAA TTTGTTGTCA  1620
      AAGTAAAAAA AAAAAAAAAA
```

5   Seq ID NO: 187 Protein sequence
     Protein Accession #: NP_000575.1

```
      1          11         21         31         41         51
      |          |          |          |          |          |
10    MTSKLAVALL AAFLISAALC EGAVLPRSAK ELRCQCIKTY SKPFHPKFIK ELRVIESGPH    60
      CANTEIIVKL SDGRELCLDP KENWVQRVVE KFLKRAENS
```

     Seq ID NO: 188 DNA sequence
     Nucleic Acid Accession #: NM_003661.1
15   Coding sequence: 1..1152

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      ATGAGTGCAC TTTTCCTTGG TGTGGGAGTG AGGGCAGAGG AAGCTGGAGC GAGGGTGCAA    60
20    CAAAACGTTC CAAGTGGGAC AGATACTGGA GATCCTCAAA GTAAGCCCCT CGGTGACTGG   120
      GCTGCTGGCA CCATGGACCC AGAGAGCAGT ATCTTTATTG AGGATGCCAT TAAGTATTTC   180
      AAGGAAAAAG TGAGCACACA GAATCTGCTA CTCCTGCTGA CTGATAATGA GGCCTGGAAC   240
      GGATTCGTGG CTGCTGCTGA ACTGCCCAGG AATGAGGCAG ATGAGCTCCG TAAAGCTCTG   300
25    GACAACCTTG CAAGACAAAT GATCATGAAA GACAAAAACT GGCACGATAA AGGCCAGCAG   360
      TACAGAAACT GGTTTCTGAA AGAGTTTCCT CGGTTGAAAA GTGAGCTTGA GGATAACATA   420
      AGAAGGCTCC GTGCCCTTGC AGATGGGGTT CAGAAGGTCC ACAAAGGCAC CACCATCGCC   480
      AATGTGGTGT CTGGCTCTCT CAGCATTTCC TCTGGCATCC TGACCCTCGT CGGCATGGGT   540
      CTGGCACCCT TCACAGAGGG AGGCAGCCTT GTACTCTTGG AACCTGGGAT GGAGTTGGGA   600
30    ATCACAGCCG CTTTGACCGG GATTACCAGC AGTACCATGG ACTACGGAAA GAAGTGGTGG   660
      ACACAAGCCC AAGCCCACGA CCTGGTCATC AAAAGCCTTG ACAAATTGAA GGAGGTGAGG   720
      GAGTTTTTGG GTGAGAACAT ATCCAACTTT CTTTCCTTAG CTGGCAATAC TTACCAACTC   780
      ATACGAGGCA TTGGGAAGGA CATCCGTGCC CTCAGACGAG CCAGAGCCAA TCTTCAGTCA   840
      GTACCGCATG CCTCAGCCTC ACGCCCCCGG GTCACTGAGC CAATCTCAGC TGAAAGCGGT   900
35    GAACAGGTGG AGAGGGTTAA TGAACCCAGC ATCCTGGAAA TGAGCAGAGG AGTCAAGCTC   960
      ACGGATGTGG CCCCTGTAAG CTTCTTTCTT GTGCTGGATG TAGTCTACCT CGTGTACGAA  1020
      TCAAAGCACT TACATGAGGG GGCAAAGTCA GAGACAGCTG AGGAGCTGAA GAAGGTGGCT  1080
      CAGGAGCTGG AGGAGAAGCT AAACATTCTC AACAATAATT ATAAGATTCT GCAGGCGGAC  1140
      CAAGAACTGT GA
```

40   Seq ID NO: 189 Protein sequence
     Protein Accession #: NP_003652.1

```
      1          11         21         31         41         51
      |          |          |          |          |          |
45    MSALFLGVGV RAEEAGARVQ QNVPSGTDTG DPQSKPLGDW AAGTMDPESS IFIEDAIKYF    60
      KEKVSTQNLL LLLTDNEAWN GFVAAAELPR NEADELRKAL DNLARQMIMK DKNWHDKGQQ   120
      YRNWFLKEFP RLKSELEDNI RRLRALADGV QKVEKGTTIA NVVSGSLSIS SGILTLVGMG   180
      LAPFTEGGSL VLLEPGMELG ITAALTGITS STMDYGKKWW TQAQAHDLVI KSLDKLKEVR   240
      EFLGENISNF LSLAGNTYQL TRGIGKDIRA LRRARANLQS VPHASASRPR VTEPISAESG   300
50    EQVERVNEPS ILEMSRGVKL TDVAPVSFFL VLDVVYLVYE SKHLHEGAKS ETAEELKKVA   360
      QELEEKLNIL NNNYKILQAD QEL
```

     Seq ID NO: 190 DNA sequence
     Nucleic Acid Accession #: NM_014452.1
55   Coding sequence: 1..1968

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      ATGGGGACCT CTCCGAGCAG CAGCACCGCC CTCGCCTCCT GCAGCCGCAT CGCCCGCCGA    60
60    GCCACAGCCA CGATGATCGC GGGCTCCCTT CTCCTGCTTG GATTCCTTAG CACCACCACA   120
      GCTCAGCCAG AACAGAAGGC CTCGAATCTC ATTGGCACAT ACCGCCATGT TGACCGTGCC   180
      ACCGGCCAGG TGCTAACCTG TGACAAGTGT CCAGCAGGAA CCTATGTCTC TGAGCATTGT   240
      ACCAACACAA GCCTGCGCGT CTGCAGCAGT TGCCCTGTGG GGACCTTTAC CAGGCATGAG   300
      AATGGCATAG AGAAATGGCCA TGACTGTAGT CAGCCATGCC CATGGCCAAT GATTGAGAAA   360
65    TTACCTTGTG CTGCCTTGAC TGACCGAGAA TGCCAGTTGC CACCTGGCAT GTTCCAGTCT   420
      AACGCTACCT GTGCCCCCCA TACGGTGTGT CCTGTGGGTT GGGGTGTGCG GAAGAAAGGG   480
      ACAGAGACTG AGGATGTGCG GTGTAAGCAG TGTGCTCGGG GTACCTTCTC AGATGTGCCT   540
      TCTAGTGTGA TGAAATGCAA AGCATACACA GACTGTCTGA GTCAGAACCT GGTGGTGATC   600
      AAGCCGGGGA CCAAGGAGAC AGACAACGTC TGTGGCACAC TCCCGTCCTT CTCCAGCTCC   660
70    ACCTCACCTT CCCCTGGCAC AGCCATCTTT CCACGCCCTG AGCACATGGA AACCCATGAA   720
      GTCCCTTCCT CCACTTATGT TCCCAAAGGC ATGAACTCAA CAGAATCCAA CTCTTCTGCC   780
      TCTGTTAGAC CAAAGGTACT GAGTAGCATC CAGGAAGGCA CAGTCCCTGA CAACACAAG   840
      TCAGCAAGCG GGAAGGAAGA CGTGAACCAA ACCCTCCCAA ACCTTCAGGT AGTCAACCAC   900
      CAGCAAGGCC CCCACCACAG ACACATCCTG AAGCTGCTGC CGTCCATGGA GGCCACTGGG   960
75    GGCGGAGAGT CCAGCACGCC CATCAAGGGC CCCAAGAGGG GACATCCTAG ACAGAACCTA  1020
      CACAAGCATT TTGACATCAA TGAGCATTTG CCCTGGATGA TTGTGCTTTT CCTGCTGCTG  1080
      GTGCTTGTGG TGATTGTGGT GTGCAGTATC CGGAAAAGCT CGAGGACTCT GAAAAAGGGG  1140
      CCCCGGCAGG ATCCCAGTGC CATTGTGGAA AAGGCAGGGC TGAAGAAATC CATGACCTCA  1200
      ACCCAGAACC GGGAGAAATG GATCTACTAC TGCAATGGCC ATGGTATCGA TATCCTGAAG  1260
80    CTTGTAGCAG CCCAAGTGGG AAGCCAGTGG AAAGATATCT ATCAGTTTCT TTGCAATGCC  1320
      AGTGAGAGGG AGGTTGCTGC TTTCTCCAAT GGGTACACAG CGGACCACGA GCGGCCCTAC  1380
      GCAGCTCTGC AGCACTGGAC CATCCGGGGC CCCGAGGCCA GCCTCGCCCA GCTAATTAGC  1440
      GCCCTGCGCC AGCACCGGAG AAACGATGTT GTGGAGAAGA TTCGTGGGCT GATGGAAGAC  1500
      ACCACCCAGC TGGAAACTGA CAAACTAGCT CTCCCGATGA GCCCCAGCCC GCTTAGCCCG  1560
```
```

```
AGCCCCATCC CCAGCCCCAA CGCGAAACTT GAGAATTCCG CTCTCCTGAC GGTGGAGCCT 1620
TCCCCACAGG ACAAGAACAA GGGCTTCTTC GTGGATGAGT CGGAGCCCCT TCTCCGCTGT 1680
GACTCTACAT CCAGCGGCTC CTCCGCGCTG AGCAGGAACG GTTCCTTTAT TACCAAAGAA 1740
AAGAAGGACA CAGTGTTGCG GCAGGTACGC CTGGACCCCT GTGACTTGCA GCCTATCTTT 1800
GATGACATGC TCCACTTTCT AAATCCTGAG GAGCTGCGGG TGATTGAAGA GATTCCCCAG 1860
GCTGAGGACA AACTAGACCG GCTATTCGAA ATTATTGGAG TCAAGAGCCA GGAAGCCAGC 1920
CAGACCCTCC TGGACTCTGT TTATAGCCAT CTTCCTGACC TGCTGTAG
```

Seq ID NO: 191 Protein sequence
Protein Accession #: NP_055267.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MGTSPSSSTA LASCSRIARR ATATMIAGSL LLLGFLSTTT AQPEQKASNL IGTYRHVDRA 60
TGQVLTCDKC PAGTYVSEHC TNTSLRVCSS CPVGTFTRHE NGIEKCHDCS QPCPWPMIEK 120
LPCAALTDRE CTCPPGMFQS NATCAPHTVC PVGWGVRKKG TETEDVRCKQ CARGTFSDVP 180
SSVMKCKAYT DCLSQNLVVI KPGTKETDNV CGTLPSFSSS TSPSPGTAIF PRPEHMETHE 240
VPSSTYVPKG MNSTESNSSA SVRPKVLSSI QEGTVPDNTS SARGKEDVNK TLPNLQVVNH 300
QQGPHHREIL KLLPSMEATG GEKSSTPIKG PKRGHPRQNL HKHFDINEHL PWMIVLFLLL 360
VLVVIVVCSI RKSSRTLKKG PRQDPSAIVE KAGLKKSMTP TQNREKWIYY CNGHGIDILK 420
LVAAQVGSQW KDIYQFLCNA SEREVAAFSN GYTADHERAY AALQHWTIRG PEASLAQLIS 480
ALRQHRRNDV VEKIRGLMED TTQLETDKLA LPMSPSPLSP SPIPSPNAKL ENSALLTVEP 540
SPQDKNKGFF VDESEPLLRC DSTSSGSSAL SRNGSFITKE KKDTVLRQVR LDPCDLQPIF 600
DDMLHFLNPE ELRVIEEIPQ AEDKLDRLFE IIGVKSQEAS QTLLDSVYSH LPDLL
```

Seq ID NO: 192 DNA sequence
Nucleic Acid Accession #: XM_044533
Coding sequence: 238..2751

```
1          11         21         31         41         51
|          |          |          |          |          |
GCTCTGCCCA AGCCGAGGCT GCGGGGCCGG CGCCGGCGGG AGGACTGCGG TGCCCCGCGG 60
AGGGGCTGAG TTTGCCAGGG CCCACTTGAC CCTGTTTCCC ACCTCCCGCC CCCCAGGTCC 120
GGAGGCGGGG GCCCCCGGGG CGACTCGGGG GCGGACCGCG GGGCGGAGCT GCCGCCCGTG 180
AGTCCGGCCG AGCCACCTGA GCCCGAGCCG CGGGACACCG TCGCTCCTGC TCTCCGAATG 240
CTGCGCACCG CGATGGGCCT GAGGAGCTGG CTCGCGCCCC CATGGGGCGC GCTGCCGCCT 300
CGGCCACCGC TGCTGCTGCT CCTGCTGCTG CTGCTCCTGC TGCTCCGACC GCCTCCGAACC 360
TGGGCGCTCA GCCCCCGGAT CAGCCTGCCT CTGGGCTCTG AAGAGCGGCC ATTCCTCAGA 420
TTCGAAGCTG AACACATCTC CAACTACACA GCCCTTCTGC TGAGCAGGGA TGGCAGGACC 480
CTGTACGTGG GTGCTCGAGA GGCCCTCTTT GCACTCAGTA GCAACCTCAG CTTCCTGCCA 540
GGCGGGGAGT ACCAGGAGCT GCTTTGGGGT GCAGACGCAG AGAAGAAACA GCAGTGCAGC 600
TTCAAGGGCA AGGACCCACA GCGCGACTGT CAAAACTACA TCAAGATCCT CCTGCCGCTC 660
AGCGGCAGTC ACCTGTTCAC CTGTGTGGCA CAGCGCCTTCA GCCCCATGTG TACCTACATC 720
AACATGGAGA ACTTCACCCT GGCAAGGGAC GAGAAGGGGA ATGTCCTCCT GGAAGATGGC 780
AAGGGCGGTT GTCCCTTCGA CCCGAATTTC AAGTCCACTG CCCTGGTGGT TGATGGCGAG 840
CTCTACACTG GAACAGTCAG CAGCTTCCAA GGGAATGACC CGGCCATCTC GCGGAGCCAA 900
AGCCTTCGCC CCACCAAGAC CGAGAGCTCC CTCAACTGGC TGCAAGACCC AGCTTTTGTG 960
GCCTCAGCCT ACATTCCTGA GAGCCTGGGC AGCTGCAAG GCGATGATGA CAAGATCTAC 1020
TTTTTCTTCA GCGAGACTGG CCAGGAATTT GAGTTCTTTG AGAACACCAT TGTGTCCCGC 1080
ATTGCCCGCA TCTGCAAGGG CGATGAGGGT GGAGAGCGGG TGCTACAGCA GCGGCTGGACC 1140
TCCTTCCTCA AGGCCCAGCT GCTGTGCTCA CGGCCCGACG ATGGCTTCCC CTTCAACGTG 1200
CTGCAGGATG TCTTCACGCT GAGCCCCAGC CCCCAGGACT GGCGTGACAC CCTTTTCTAT 1260
GGGGTCTTCA CTTCCCAGTG GCACAGGGGA ACTACAGAAG GCTCTGCCGT CTGTGTCTTC 1320
ACAATGAAGG ATGTGCAGAG AGTCTTCAGC GGCCTCTACA AGGAGGTGAA CCGTGAGACA 1380
CAGCAGTGGT ACACCGTGAC CCACCCGGTG CCCACACCCC GGCCTGGAGC GTGCATCACC 1440
AACAGTGCCC GGGAAAGGAA GATCAACTCA TCCCTGCAGC TCCCAGACCG GCTGCTGAAC 1500
TTCCTCAAGG ACCACTTCCT GATGGACGGG CAGGTCCGAA GCCGCATGCT GCTGCTGCAG 1560
CCCCAGGCTC GCTACCAGCG CGTGGCTGTA CACCGCGTCC CTGGCCTGCA CCACACCTAC 1620
GATGTCCTCT TCCTGGGCAC TGGTGACGGC CGGCTCCACA AGGCAGTGAG CGTGGGCCCC 1680
CGGGTGCACA TCATTGAGGA GCTGCAGATC TTCTCATCGG GACAGCCCGT GCAGAATCTG 1740
CTCCTGGACA CCCACAGGGG GCTGCTGTAT GGGGCCTCAC ACTCGGGCGT AGTCCAGGTG 1800
CCCATGGCCA ACTGCAGCCT GTACAGGAGC TGTGGGGACT CGGGGACCCC CCGGGACCCC 1860
TACTGTGCTT GGAGCGGGCTC CAGCTGCAAG CACGTCAGCC TCTACCAGCC TCAGCTGGCC 1920
ACCAGGCCGT GGATCCAGGA CATCGAGGGA GCCAGCGCCA AGGACCTTTG CAGCGCGTCT 1980
TCGGTTGTGT CCCCGTCTTT TGTACCAACA GGGGAGAAGC CATGTGGACA AGTCCAGTTC 2040
CAGCCCAACA CAGTGAACAC TTTGGCCTGC CGGCTCCTCT CCAACCTGGC GACCCGACTC 2100
TGGCTACGCA ACGGGGCCCC CGTCAATGCC TCGGCCTCCT GCCACGTGCT ACCCACTGGG 2160
GACCTGCTGC TGGTGGGCAC CCAACAGCTG GGGGAGTTCC AGTGCTGGTC ACTAGAGGAG 2220
GGCTTCCAGC AGCTGGTAGC CAGCTACTGC CCAGAGGTGG TGGAGGACGG GGTGGCAGAC 2280
CAAACAGATG AGGGTGGCAG TGTACCCGTC ATTATCAGCA CATCGCGTGT GAGTGCACCA 2340
GCTGGTGGCA AGGCCAGCTG GGGTGCAGAC AGGTCCTACT GGAAGGGAGTT CCTGGTGATG 2400
TGCACGCTGT TTGTGCTGGC CGTGCTGCTC CCAGTTTTGT TCTTGCTCTA CCGGCACCGG 2460
AACAGCATGA AAGTCTTCCT GAAGCAGGGG GAATGTGCCA GCGTGCACCC CAAGACCTGC 2520
CCTGTGGTGC TGCCCCCTGA GACCCGCCCA CTCAACGGCC TAGGGCCCCC TAGCACCCCG 2580
CTCGATCACC GAGGGGTACC AGTCCCTGTCA GACAGCCCCG CGGGGTCCCG AGTCTTCACT 2640
GAGTCAGAGA AGAGGCCACT CAGCATCCAA GACAGCTTCG TGGAGGTATC CCCAGTGTGC 2700
CCCCGGCCCC GGGTCCGCCT TGGCTCGGAG ATCCGTGACT CTGTGGTGTG AGAGCTGACT 2760
TCCAGAGGAC GCTGCCCTGG CTTCAGGGGC TGTGAATGCT CAACTGGACC 2820
TCCCCTCCGC TCTGCTCTTC GTGGAACACG ACCGTGGTGC CCGGCCCTTG GGAGCCTTGG 2880
GGCCAGCTGG CCTGCTGCTC TCCAGTCAAG TAGCCAAGCT CCTACCACCC AGACACCCAA 2940
ACAGCCGTGG CCCCAGAGGT CCTGGCCAAA TATGGGGGAC TGCCTAGGTT GGTGGAACCG 3000
TGCTCCTTAT GTAAACTGAG CCCTTTGTTT AAAAAACAAT TCCAAATGTG AAACTAGAAT 3060
GAGAGGGAAG AGATAGCATG GCATGCAGCA CACACGGCTG CTCCAGTTCA TGGCCTCCCA 3120
GGGGTGCTGG GGATGCATCC AAAGTGGTTG TCTGAGACAG AGTTGGAAAC CCTCACCAAC 3180
```

```
TGGCCTCTTC ACCTTCCACA TTATCCCGCT GCCACCGGCT GCCCTGTCTC ACTGCAGATT  3240
CAGGACCAGC TTGGGCTGCG TGCGTTCTGC CTTGCCAGTC AGCCGAGGAT GTAGTTGTTG  3300
CTGCCGTCGT CCCACCACCT CAGGGACCAG AGGGCTAGGT TGGCACTGCG GCCCTCACCA  3360
GGTCCTGGGC TCGGACCCAA CTCCTGGACC TTTCCAGCCT GTATCAGGCT GTGGCCACAC  3420
GAGAGGACAG CGCGAGCTCA GGAGAGATTT CGTCACAATG TACGCCTTTC CCTCAGAATT  3480
CAGGGAAGAG ACTGTCGCCT GCCTTCCTCC GTTGTTGCGT GAGACCCCGT GTGCCCCTTC  3540
CCACCATATC CACCCTCGCT CCATCTTTGA ACTCAAACAC GAGGAACTAA CTGCACCCTG  3600
GTCCTCTCCC CAGTCCCCAG TTCACCCTCC ATCCCTCACC TTCCTCCACT CTAAGGGATA  3660
TCAACACTGC CCAGCACAGG GGCCCTGAAT TTATGTGGTT TTTATACATT TTTTAATAAG  3720
ATGCACTTTA TGTCATTTTT TAATAAAGTC TGAAGAATTA CTGTTT
```

Seq ID NO: 193 Protein sequence
Protein Accession #: XP_044533.3

```
1          11         21         31         41         51
|          |          |          |          |          |
MLRTAMGLRS WLAAPWGALP PRPPLLLLLL LLLLLQPPPP TWALSPRISL PLGSEERPFL   60
RPEAEHISNY TALLLSRDGR TLYVGARBAL FALSSNLSPL PGGEYQELLW GADAEKKQQC  120
SFKGKDPQRD CQNYIKILLP LSGSHLPTCG TAAFSPMCTY INMENFTLAR DEKGNVLLED  180
GKGRCPFDPN FKSTALVVDG ELYTGTVSSF QGNDPAISRS QSLRPTKTES SLNWLQDPAF  240
VASAYIPESL GSLQGDDDKI YFFFSETGQE FEFFENTIVS RIARICKGDE GGERVLQQRW  300
TSFLKAQLLC SRPDDGFPFN VLQDVFTLSP SPQDWRDTLF YGVFTSQWER GTTEGSAVCV  360
FTMKDVQRVF SGLYKEVMRE TQQNYTVTHP VPTPRPQACI TNSARERKIN SSLQLPDRVL  420
NFLKDHFLMD GQVRSRMLLL QPQARYQRVA VHRVPGLHHT YDVLPLGTGD GRLHKAVSVG  480
PRVHIIEELQ IFSSGQPVQN LLLDTHRGLL YAASHSGVVQ VPMANCSLYR SCGDCLLARD  540
PYCAWSGSSC KHVSLYQPQL ATRPWIQDIE GASAKDLCSA SSVVSPSFVP TGEKPCEQVQ  600
FQPNTVNTLA CPLLSNLATR LWLRNGAFVN ASASCHVLPT GDLLLVGTQQ LGEFQCWSLE  660
EGFQQLVASY CFEVVEDGVA DQTDEGGSVP VIISTSRVSA PAGGKASWGA DRSYWKEFLV  720
MCTLFVLAVL LPVLFLLYRH RNSMKVFLKQ GECASVHPKT CPVVLPPETR PLNGLGPPST  780
PLDHRGYQSL SDSPPGSRVF TESEKRPLSI QDSPVEVSPV CPRPRVRLGS EIRDSVV
```

Seq ID NO: 194 DNA sequence
Nucleic Acid Accession #: NM_022819.1
Coding sequence: 1..635

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGCAGATG GGGCAAAGGC CAACCCCAAA GGGTTCAAAA AGAAGGTGCT GGATAGATGC   60
TTCTCTGGGT GGAGGGGCCC ACGCTTCGGG GCCTCCTGTC CTTCAAGAAC CTCCAGGTCT  120
AGCCTGGGTA TGAAGAAGTT CTTCACCGTG GCCATCCTTG CTGGCAGCGT TCTGTCCACA  180
GCTCACGGCA GCCTGCTCAA CCTGAAGGCC ATGGTGGAGG CCGTCACAGG GAGGAGCGCC  240
ATCCTGTCCT TCGTGGGCTA CGGTTGCTAC TGTGGGCTGG GGGCCGTGG CCAGCCCAAG  300
GATGAGGTGG ACTGGTGCTG CCACGCCCAC GACTGCTGCT ACCAGGAACT CTTTGACCAA  360
GGCTGTCACC CCTATGTGGA CCACTATGAT CACACCATCG AGAACAACAC TGAGATAGTC  420
TGCAGTGACC TCAACAAGAC AGAGTGTGAC AAGCAGACAT GCATGTGTGA CAAGAACATG  480
GTTCTGTGCC TCATGAACCA GACGTACCGA GAGGAGTACC GTGGCTTCCT CAATGTCTAC  540
TGCCAGGGCC CCACGCCCAA CTGCAGCATC TATGAACTGC CCCCTGAGGA GGTCACCTGC  600
AGTCACCAAT CCCCAGCGCC CCCCGCCCCT CCCTAG
```

Seq ID NO: 195 Protein sequence
Protein Accession #: NP_073730

```
1          11         21         31         41         51
|          |          |          |          |          |
MADGAKANPK GFKKKVLDRC FSGWRGPRFG ASCPSRTSRS SLGMKKFPTV AILAGSVLST   60
AEGSLLNLKA MVEAVTGRSA ILSFVGYGCY CGLGGRGQPK DEVDWCCHAH DCCYQELFDQ  120
GCHPYVDHYD HTIENNTEIV CSDLNKTECD KQTCMCDKNM VLCLMNQTYR EEYRGFLNVY  180
CQGPTPNCSI YEPPPEEVTC SHQSPAPPAP P
```

Seq ID NO: 196 DNA sequence
Nucleic Acid Accession #: XM_028196.1
Coding sequence: 1315..1791

```
1          11         21         31         41         51
|          |          |          |          |          |
GGCATTGATG CTGTGTGCGC GTGCGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT   60
GTGTGTGTGT GTCTGGAGTC ATGGCAGGGT CCCTTTCTGT CTGTCTCCTT GCTCTGCCCC  120
AGACTGGGGG GCTGCAGAGG TGAGGGTATC TGGCCTCAAC AGCTGCTTAT TCCCGATGGG  180
ATGGCCTGGG CTGGGCCCCT GAGGCCAGGC TGACTTGGAC ATGGCAAGAG GGGTCCCAGG  240
CTCTTGTGGG CAAAGCAGGG GAGGCGCCAA TGTGGAGGAA CAGAGTCTCC TGGCTGGCTG  300
CTGCTCTCTG GAGCGGGTGG AGTCAGGGAA GAGCTGAGCT GGGGAGTCAC CCTGGGCCTG  360
GGGTCACCGT AGGCCCCATG TAGCACCCTG GTTCCCCTGC CTGTAGGTGA CAGGAGCCAG  420
CCCAGCCAGG TGTGCTCCCT CCCCAGGCCC TAGGCAGGCG GGTACAGGGG CCAGCAGCTG  480
CGCCCGCCCC ACCTTCCTTC CCACCCACAT GCCGAAGGGT GGCCAGGCAG GCAGGTGGAC  540
GAGTCCAGGC AGCGGCTGAG TCAGTGTGTG TGGAATGTTC TGGCCGCTCC CAGCTGCACC  600
CTGCCCCTAC CTGCCACCAC CTCACCTTCA TCCTCAGGCG CTGCGGCCCT GAGCCCCTGC  660
CAGGAATGCA CCTTTAGCCC AGGCCTGCTC AGTGAGCTCC GCCGACAGCC AGCCCTGCTC  720
CTCCCGCCAT GACCCTGCAG ACCCCTCTGG GCTTCCAAGT TCCTGGGGGC TGCAGTGAAC  780
ATGCTCCACC TGCATGGCTG GCAAACCATG GTGGGCCCCA GCTGTGGTGC GTGCTGGGGT  840
AGAGGCAAGG AAGTGATGGG ACCGCAGAGA TGAGACCCCC AGGGATGAGA TGGGACCCCC  900
AGGCAGGGCC CAGGGTCCAG GGCCCAGGAG AGAGAAGCAG GGAGGGAGAG AGCTTCCTGG  960
TGGAGGACGC ATCCTACAGT GGGGGCAAGG GTGCTCTGAG GTCCGGTGAA GGCAGGGACT 1020
AGGCTGCCCA GGCCCTGCCT GCTTGGCTGG GGCTGGGGGC TGCTGGGAGG TGGCTGGGAG 1080
GCTGGGCCTG GGCAGCTAAG CTGGAGCTTT GGCCAGGGTC CAGAGCCTCC CTCCCTTCAG 1140
```

```
CTTCCTGCTG CACAGAACCC TCGCCCCTGG CCACCCCGTG CTGCCTCCTT GCCCTGGCAG   1200
ACCCAGCACT GGCTGCTGCT AGTCAGATGG GGTAGCGGGC AGGGGCCGGA GGGGCCCACCC   1260
TCCCAGCTGA CCCAGCCTCC TGGGCCGGCTT CTTCCAAACC AGCAGGGTAG AAAGATGGGG   1320
CACCCACCAG TCTCTCCCAG TGCCCCGGCC CCAGCTGGCA CCACAGCTAT ACCTGGGCTT   1380
ATTCCAGACC TTGTCGCCGG GACCCCCTGT GAGTTGTGGG ATTCCCAAGA GGGGTGTGGG   1440
GATAACCCAG CCAAGTGGGG GCTGCAGCTG TCCACAGATG CACTCAGCGT GGCCTCTACC   1500
CCAGGGCCCC GCTGGGCTCT CATTGCCGGC GCCCTTGCCG CGGGCGTCCT CCTCGTCTCC   1560
TGCCTCCTCT GTGCTGCCTG CTGCTGCTGC CGCCGCCACA GGAAGAAGCC CAGGGACAAG   1620
GAGTCCGTGG GTCTGGGCAG TGCCCGCGGC ACCACCACCA CCCACCTGGT GAGGAGCGGC   1680
TCCTTGCTCA CTCAGTCCAG AGAGGGCTTG AAATCCAGGC TCCAGAGCCC AGGGCAGCGA   1740
GGCGAGTTCA GCCCCAGGGA TGGTTTAACC CCCACAGAGG CAGGGCGTTG AGGACCTTCC   1800
TGGCAGGGAA AGTGGGTGAA CAGAGGTGAG AAGGAGGCCA TGCAACAGGG GCTGCCCCAT   1860
GGGCCCGAGG GAGCCACAGC GGGTTCTTGA GGAAGGCAGG GGGTACCCCA GATGCCACGT   1920
TTTGGGTGGG TTTGGCCGGT CTCACAGAGC GAAGCCGACG ATTTGTGCCT GTTGGGTGGC   1980
CTGGCCTGGA GGCGGGGGGT CTTGACCCAT GTCATGCAAG GGCTGCCCGG GAGCCCAGGG   2040
CTCTGATGAG GCATGATGTC AGCACCACCT GCCCCTTGTC CCAACTCACT CCAGGTGCAA   2100
CCTGATGTGG ATGGCCTGGA GTCCAGCCCG GGGGATGCTC AGCAATGGGG GTGCCTGCAG   2160
CTCTCCCTGG AGTTCGACTT TGGAAGCCAG GAGGTGAAGG GCCCCGCTGC GCAGGACCAG   2220
CGGTTCTGCG AGTTTCCGGA AAGGGTGACG GGGGAAGGGC AGACCCCATG CCCTGGGTGG   2280
TGGGGAGCTG ACAGGGCAGG GGCCCTTGGC TGAGCCCACC CCGCTGGCTC CCAGATCAGG   2340
GTGGGCCTGA GGCAGGCAGC CGACCTGAGG CCTGGGGGCA CCGTGGACCC CTATGCCCGG   2400
GTCAGCGTCT CCACCCAGGC CGGACACAGA CATGAGACAA AAGTGCACCG AGGGCACGCTC   2460
TGCCCCGTGT TTGACGAGAC CTGCTGCTTC CACGTGAGTC AGGGATGGTC GGCTGGGTGG   2520
GCCTGGACGG CTGGATGGGC CTGGGCTGGG TGGGCCTGGG CAGCTGGGTG GGCCTGGGCA   2580
GCTGGGTGGG CCTGAGCTAG GGCAGCAGGG CCTGGCTCAC GCCGCTGCCT CAGATCCCGC   2640
AGGCGGAGCT GCCCAGGGGCC ACCCTGCAGG TGCAGCTTTT CAACTTCAAG CGCTTCTCGG   2700
GGCATGAGCC CCTGGGTGAG CTCCGTCTGC CACTGGGCAC CGTGGATCTG CAGCATGTTC   2760
TGGAGCACTG GTACCTGCTG GGCCCGCCGG CTGCCACTCA GGTGAGGTGC TGGTCACCAG   2820
GCCCACAGCCC AAGGCAGAGC TGGCAGGGAC CCTGCCCTAT GGGCCATCGG AAAGACAGGC   2880
CTGATGGGCA GCATTTTCGG GGGTCTGAGC CCCAACTCGG CCAGAATCAC CCTCCCGGGC   2940
TGAAGCCCCT CTTGCTGCCC ACAGCCCGAG CAGGTCGGGG AGCTGTGCTT CTCTCTCCGG   3000
TACGTGCCCA GCTCAGGCCG GCTGACCGTG GTGGTGCTGG AGGCTCGAGG CCTGCGTCCA   3060
GGACTTGCAG AGCCCTACGT GAAGGTCCAG CTCATGCTGA ACCAGAGGAA GTGGAAGAAG   3120
AGAAAGACAG CCACCAAAAA GGGCACGGCG GCCCCCTACT TCAATGAGGC CTTCACCTTC   3180
CTGGTGCCCT TCAGCCAGGT CCAGAATGTG GACCTGGTGC TGGCTGTCTG GGACCGCAGC   3240
CTGCCGCTCC GAACTGAGCC CGTAGGCAAG GTGCACCTGG GTGCCCGGGC CTCGGGGCAG   3300
CCCCTGCAGC ACTGGGCAGA CATGCTGGCC CACGCCCGGC GGCCCATTGC CCAGCGGCAC   3360
CCCCTGCCCG CAGCCAGGGA GGTGGACCGC ATGCTGGCCC TGCAGCCCCG CCTTCGCCTG   3420
CGCCTGCCCT TGCCCCACTC CTGAATGCAC CACCATGCCTC TGTCTCCCCG CTGAGCCCAG   3480
GCACTTGCCC AGGCCGCCCT GCAGGACCAC TGCAATAAAC GCCTTCTCCT GCC
```

Seq ID NO: 197 Protein sequence
Protein Accession #:   XP_028196.1

```
 1          11          21          31          41          51
 |          |           |           |           |           |
MGHPPVSPSA PAPAGTTAIP GLIPDLVAGT PCELWDSQBG CGDNPAKWGL QLSTDALSLA    60
STPGPRWALI AGALAAGVLL VSCLLCAACC CCRRHRKKPR DKESVGLGSA RGTTTTHLVR   120
SGSLLTQSRE GLKSRLQSPG QRGEFSPRDG LTPTEAGR
```

Seq ID NO: 198 DNA sequence
Nucleic Acid Accession #: NM_000612.2
Coding sequence: 553..1095

```
 1          11          21          31          41          51
 |          |           |           |           |           |
TTCTCCCGCA ACCTTCCCTT CGCTCCCTCC CGTCCCCCCC AGCTCCTAGC CTCCGACTCC    60
CTCCCCCCCT CACGCCCGCC CTCTCGCCTT CGCCGAACCA AAGTGGATTA ATTACACGCT   120
TTCTGTTTCT CTCCGTGCTG TTCTCTCCCG CTGTGCGCCT GCCCGCCTCT CGCTGTCCTC   180
TCTCCCCTC GCCCTCTCTT CGGCCCCCCC CTTTCACGTT CACTCTGTCT CTCCCACTAT   240
CTCTGCCCCC CTCTATCCTT GATACAACAG CTGACCTCAT TTCCCGATAC CTTTTCCCCC   300
CCGAAAAGTA CAACATCTGG CCCGCCCCAG CCCGAAGACA GCCCGTCCTC CCTGGACAAT   360
CAGACGAATT CTCCCCCCCC CCCAAAAAAA AAAAGCCATC CCCCCGCTCT GCCCCGTCGC   420
ACATTCGGCC CCCGCGACTC GGCCAGAGCG GCGCTGGCAG AGGAGTGTCC GGCAGGAGGG   480
CCAACGCCCG CTGTTCGGTT TGCGACACGC AGCAGGGAGG TGGGCGGCAG CGTCGCCGGC   540
TTCCAGACAC CAATGGGAAT CCCAATGGGG AAGTCGATGC TGGTGCTTCT CACCTTCTTG   600
GCCTTCGCCT CGTGCTGCAT TGCTGCTTAC CGCCCCAGTG AGACCCTGTG CGGCGGGGAG   660
CTGGTGGACA CCCTCCAGTT CGTCTGTGGG GACCGCGGCT TCTACTTCAG CAGGCCGGCA   720
AGCCGTGTGA GCCGTCGCAG CCGTGGCATC GTTGAGGAGT GCTGTTTCCG CAGCTGTGAC   780
CTGGCCCTCC TGGAGACGTA CTGTGCTACC CCCGCCAAGT CCGAGAGGGA CGTGTCGACC   840
CCTCCGACCG TGCTTCCGGA CAACTTCCCC AGATACCCCG TGGGCAAGTT CTTCCAATAT   900
GACACCTGGA AGCAGTCCAC CCAGCGCCTG CGCAGGGGCC TGCCTGCCCT CCTGCGTGCC   960
CGCCGGGGTC ACGTGCTCGC CAAGGAGCTC GAGGCGTTCA GGGAGGCCAA ACGTCACCGT  1020
CCCCTGATTG CTCTACCCAC CCAAGACCCC GCCCACGGGG GCGCCCCCCC AGAGATGGCC  1080
AGCAATCGGA AGTGAGCAAA ACTGCCGCAA GTCTGCAGCC CGGCGCCACC ATCCTGCAGC  1140
CTCCTCCTGA CCACGGACGT TTCCATCAGG TTCCATCCG AAAATCTCTC GGTTCCACGT  1200
CCCCCTGGGG CTTCTCCTGA CCCAGTCCCC GTGCCCCGCC TCCCCGAAAC AGGCTACTCT  1260
CCTCGGCCCC CTCCATCGGG CTGAGGAAGC ACAGCAGCAT CTTCAAACAT GTACAAAATC  1320
GATTGGCTTT AAACACCCTT CACATACCCT CCCCCC
```

Seq ID NO: 199 Protein sequence
Protein Accession #: NP_000603.1

```
 1          11          21          31          41          51
```

```
     AAAGTTCCCA GCCAATAGAC AGCATGAATC AAGGAACTTG CATTATATGT GCTCTTGAAT   1680
     CTGTTGTCTC CATGGACCAT TCCTCGGAGT AGTGGTGAGA TGGCCTTGGG TTGCCCTTGG   1740
 5   CTTCTCCTCC CTCTACTCAG CCTTAAAAAG GGCTTCTTGG AACTTTACCA GCAGCCTCAG   1800
     CTTTACAAAT GCCTTGGTAT GTACCTCTGG CAAATGCCCT GGTATGTACC TCTGGCAAAT   1860
     GCCCCACCTT GGTGATGTTG CAACCTTTCC TTCTGCTAGG GTGTACACCT AGCCTGTGCA   1920
     GGTGTCAGCC CTGCTAGGGA GTCACTGTAC ACACAAACTC TACTGGAATT CCTGCCAACA   1980
     TCTGTCACCC TGCAGCTCCT TTACAGTTCA ATCCAATGAT AGAAACCATC CCTTCCCTTT   2040
     CTCCCTTGGC TGTTCACCCA GCCATTCCCT GAAGGCCTTA CCAACAGGAA TATCCAAGAA   2100
 10  GCTGTTGTCC TCTCTCGAAC CCTGACCAGA TCATCAGCCA CTGAGGCCAG TGGAATTTCC   2160
     CCAGGCCTTG TTAAAAAAAA AAAAAAAAAA                                    2190
```

Seq ID NO: C66 DNA Sequence
Nucleic Acid Accession #: NM_014459.2
Coding sequence: 738..3407

```
    1         11         21         31         41         51
    |         |          |          |          |          |
    GTAGATGCAG TCCGCCGCCG CCGCTGCCTC AGCCAGCAAT GCAAGATTAG ATCTCTAAAT   60
    GCAGCAAAAC ACTGCCTGAA AACAGACCGG CCCGCGCAGC AAGCAGACAT TTCACGGTGC   120
 20 GCTGGGGAAG CTTCAAAATA TATCTGTGAC TCTGTCTTCG TTGCTCTTCA TCCCCATCAA   180
    TTTCATCACG GGAGGCGAGC AGCAAGTAAG AATTTCACTT TCGGATCTGC CTAGAGACAC   240
    ACCTCCCTGC TCCCTCCCCC ACTCGATGTG AAGAGTATTC CGGAGTCTCC GGGCGGGAGT   300
    AGATTTGCAG CACCCTAGCG GGAGCGAGGA AAACCTACTG ATTCTTTAGC TCATTATCAT   360
    CTCTCCCAGA CGAGATTTCC TTCTTATCGC CTGCCTCATC GCTCAAGTTT GAGCCTCCCG   420
 25 AAGTCCGGGC GGGAGAGACG AAACCCCTGG CTCACCCCCA GCCGCAGGAA GCCACCGGCT   480
    TGCTCCAAGC CCCTGCAGCT CTGCTGCACC GCAGCTTCTC ACCCAGTGCG GATGCTGTAG   540
    ATCAACAGGT TCAGGGAACT TGAGCAGAAT AAGGAGAGCA CCCGGGTGC CGCAGCTCGG    600
    GTGCAGAGGG AAAAAAGGAC CCATAGACTT GTGGCTCGCG TCGCGCGCGC ACGCTGCGCC   660
    AGGGCCCCAG GCTGGCGCGC ACTCCCTCTC TGGCTCCTCC AGTCCGATTG CTCCTGCCCC   720
 30 CACCTTACAG GTCTGGGATG TACCTTTCCA TCTGTTGCTG CTTTCTTCTA TGGGCCCCTG   780
    CCCTCACTCT CAAGAACCTC AACTACTCCG TGCCGGAGGA GCAAGGGGCC GGCACGGTGA   840
    TCGGGAACAT CGGCAGGGAT GCTCGACTGC AGCCTGGGCT TCCGCCTGCA GAGCGCGGCG   900
    GGGGAGGGCG CAGCAAGTCG GGTAGCTACC GGGTGCTGGA GAACTCCGCA CCGCACCTGC   960
 35 TGGACGTGGA CGCAGACAGC GGGCTCCTCT ACACCAAGCA GCGCATCGAC CGCGAGTCCC   1020
    TGTGCCGCCA CAATGCCAAG TGCCAGCTGT CCCTCGAGGT GTTCGCCAAC GACAAGGAGA   1080
    TCTGCATGAT CAAGGTAGAG ATCCAGGACA TCAACGACAA CGCGCCCTCC TTCTCCTCGG   1140
    ACCAGATCGA AATGGACATC TCGGAGAACG CTGCTCCGGG CACCCGCTTC CCCCTCACCA   1200
    GCGCACATGA CCCCGACGCC GGCGAGAATG GGCTCCGCAC CTACCTGCTC ACGCGCGACG   1260
    ATCACGGCCT CTTTGGACTG GACGTTAAGT CCCGCGGCGA CGGCACCAAG TTCCCAGAAC   1320
 40 TGGTCATCCA GAAGGCTCTG GACCGCGAGC AACAGAATCA CCATACGCTC GTGCTGACTG   1380
    CCCTGGACGG TGGCGAGCCT CCACGTTCCG CCACCGTACA GATCAACGTG AAGGTGATTG   1440
    ACTCCAACGA CAACAGCCCG GTCTTCGAGG CGCCATCCTA CTTGGTGGAA CTGCCCGAGA   1500
    ACGCTCCGCT GGGTACAGTG GTCATCGATC TGAACGCCAC CGACGCCGAT GAAGGTCCCA   1560
    ATGGTGAAGT GCTCTACTCT TTCAGCAGCT ACGTGCCTGA CCGCGTGCGG GAGCTCTTCT   1620
 45 CCATCGACCC CAAGACCGGC CTAATCCGTG TGAAGGGCAA TCTGGACTAT GAGGAAAACG   1680
    GGATGCTGGA GATTGACGTG CAGGCCCGAG ACCTGGGGCC TAACCCTATC CCAGCCCACT   1740
    GCAAAGTCAC GGTCAAGCTC ATCGACCGCA ACGACAATGC GCCGTCCATC GGTTTCGTCT   1800
    CCGTGCGCCA GGGGGGCGCTG AGCGAGGCCG CCCCTCCCGG CACCGTCATC GCCCTGGTGC   1860
    GGGTCACTGA CCGGGACTCT GGCAAGAACG GACAGCTGCA GTGTCGGGTC CTAGGCGGAG   1920
 50 GAGGGACGGG CGGCGGCGGG GGCCTGGGCG GGCCGGGGGG TTCCGTCCCC TTCAAGCTTG   1980
    AGGAGAACTA CGACAACTTC TACACGGTGG TGACTGACCG CCCGCTGGAC CGCGAGACAC   2040
    AAGACGAGTA CAACGTGACC ATCGTGGCGC GGGACGGGGG CTCTCCTCCC CTCAACTCCA   2100
    CCAAGTCGTT CGCGATCAAG ATTCTAGACG AGAACGACAA CCCGCCTCGG TTCACCAAAG   2160
 55 GGCTCTACGT GCTTCAGGTG CACGAGAACA ACATCCCGGG AGAGTACCTG GGCTCTGTGC   2220
    TCGCCCAGGA TCCCGACCTG GGCCAGAACG GCACCGTATC CTACTCTATC CTGCCCCAGT   2280
    ACATCGGCGA CGTGTCTATC TACACCCTATG TGTCTGTGAA TCCCACGAAC GGGGCCATCT   2340
    ACGCCCTGCG CTCCTTTAAC TTCGAGCAGA CCAAGGCTTT TGAGTTCAAG GTGCTGCTA    2400
    AGGACTCGGG GGCGCCCGCG CACTTGGAGA GCAACGCCAC GGTGAGGGTG ACAGTGCTAG   2460
 60 ACGTGAATGA CAACGCGCCA GTGATCGTGC TCCCCACGCT GCAGAACGAC ACCGCGGGACG   2520
    TGCAGGTGCC GCGCAACGCT GGCCTGGGGCT ATCTGGTGAG CACTGTGCGC GCCCTAGACA   2580
    GCGACTTCGG CGAGAGCGGG CGTCTCACCT ACGAGATCGT GGACGGCAAC GACGACCACC   2640
    TGTTTGAGAT CGACCGGTCC AGCGGCGAGA TCCGCACGCT GCACCCTTTC TGGGAGGACG   2700
    TGACGCCCGT GGTGGAGCTG GTGGTGAAGG TGACCGACCA CGGCCAAGCCT ACCCTGTCCG   2760
 65 CAGTGGCCAA GCTCATCATC CGCTCGGTGA GCGGATCCCT TCCCGAGGGG GTACCACGGG   2820
    TGAATGGCGA GCAGCACCAC TGGGACATGT CGCTGCCCGT CATCGTGACT CTGAGCACTA   2880
    TCTCCATCAT CCTCCTAGCG GCCATGATCA CCATCGCCGT CAAGTGCAAG CGCGAGAACA   2940
    AGGAGATCCG CACTTACAAC TGCCGCATCG CCGAGTACAG CCACCCGCAG CTGGGTGGGG   3000
    GCAAGGGCAA GAAGAAGAAG ATCAACAAAA ATGATATCAT GCTGGTGCAG AGCGAAGTGG   3060
 70 AGGAGAGGAA CGCCATGAAC GTCATGAACG TGGTGAGCAG CCCCTCCCTG GCCACCTCCC   3120
    CCATGTACTT CGACTACCAG ACCCGCCTGC CCCTCAGCTC GCCCCGGTCG GAGGTGATGT   3180
    ATCTCAAACC GGCCTCCAAC AACCTGACTG TCCCTCAGGG GCACGCGGGC TGCCACACCA   3240
    GCTTCACCGG ACAAGGGACT AATGCAAGCG AGACCCCTGC CACTCGGATG TCCATAATTC   3300
 75 AGACAGACAA TTTTCCCGCA GAGCCCAATT ACATGGGCAG CAGGCAGCAG TTTGTTCAAA   3360
    GTATTTCAGT AGCTCCACGT TTAAGGACCC AGAAAGAGCC AGCCTGAGAG ACAGTGGGCA   3420
    CGGGGACAGT GATCAGGCTG ACAGTGACCA AGACACTAAC AAAAGGCTCCT GCTGTGACAT   3480
    GTCTGTTAGG GAGGCACTCA AGATGAAAAC TACTTCAACT AAAAGCCAAC CACTTGAACA   3540
    AGAACCAGAA GAGTGTGTTA ATTGCACAGA TGAATGCCGA GTGCTTGGTC ATTCTGACAG   3600
    GTGCTGGATG CCACAGTTCC CTGCAGCCAA TCAGGCTGAA AATGCAGATT ACCGCACAAA   3660
    TCTCTTTGTA CCTACAGTTG AAGCTAATGT TGAGACTGAG ACTTACGAAA CTGTGAATCC   3720
 80 CACTGGGAAA AAGACTTTTT GTACATTTGG AAAAGACAAG CGAGAGCACA CTATTCTCAT   3780
    TGCCAACGTT AAACCTTATT TAAAAGCCAA ACGTGCCCTG AGCCCTCTCC TCCAAGAGGT   3840
    CCCCTCAGCA TCAAGCAGCC CAACCAAGGC GTGCATCGAG CCTTGCACCT CAACAAAAGG   3900
    CTCCCTGGAT GGCTGTGAAG CAAAACCAGG AGCCCTGGCT GAAGCAAGCA GTCAGTACTT   3960
    GCCCACTGAC AGTCAATATC TGTCACCTAG TAAGCAACCA AGAGACCCTC CCTTCATGGC   4020
```

1237

```
                    TCAGCCCACC CTTGGGTCCG GGCCAACTCT CGGAGGGTGC TGCCTCCCTC TGCCCTTCAA 1080
                    TCTGTCGCCT GATGGTCCCT GTCATTCACT CGGGTGCGTG TGTTTGTATG TCTGTGTATG 1140
                    TATAGGGGAA AGAAGGGATC CCTAACTGTT CCCTTATCTG TTTTCTACCT CCTCCTTTGT 1200
     5              TTAATAAAGG CTGAAGCTTT TTGT

                    Seq ID NO: 203 Protein sequence
                    Protein Accession #: NP_004208

     10             1          11         21         31         41         51
                    |          |          |          |          |          |
                    MAQKENSYPW PYGRQTAPSG LSTLPQRVLR KEPVTPSALV LMSRSNVQPT AAPGQKVMEN    60
                    SSGTPDILTR HFTIDDPEIG RPLGKGKFGN VYLAREKKSH FIVALKVLFK SQIEKEGVEH   120
                    QLRREIEIQA HLHHPNILRL YNYFYDRRRI YLILEYAPRG ELYKELQKSC TFDEQRTATI   180
     15             MEELADALMY CHGKKVIHRD IKPENLLLGL KGELKIADFG WSVHAPSLRR KTMCGTLDYL   240
                    PPEMIEGRMH NEKVDLWCIG VLCYELLVGN PPFESASHNE TYRRIVKVDL KFPASVPTGA   300
                    QDLISKLLRH NPSERLPLAQ VSAHPWVRAN SRRVLPPSAL QSVA

                    Seq ID NO: 204 DNA sequence
                    Nucleic Acid Accession #: AK055663
     20             Coding sequence: 38..1423


                    1          11         21         31         41         51
     25             |          |          |          |          |          |
                    AGAACGGCTT CCGGCGGGAG CTGTGCAGCT CCTTATCATG GGGACAATTC ATCTCTTTCG    60
                    AAAACCACAA AGATCCTTTT TTGGCAAGTT GTTACGGGAA TTTAGACTTG TAGCAGCTGA   120
                    CCGAAGGTCC TGGAAGATAC TGCTCTTTGG TGTAATAAAC TTGATATGTA CTGGCTTCCT   180
                    GCTTATGTGG TGCAGTTCTA CTAATAGTAT AGCTTTAACT GCCTATACTT ACCTGACCAT   240
     30             TTTTGATCTT TTTAGTTTAA TGACATGTTT AATAAGTTAC TGGGTAACAT TGAGGAAACC   300
                    TAGCCCTGTC TATTCATTTG GGTTTGAAAG ATTAGAAGTC CTGGCTGTAT TTGCCTCCAC   360
                    AGTCTTGGCA CAGTTGGGAG CTCTCTTTAT ATTAAAAGAA AGTGCAGAAC GCTTTTTGGA   420
                    ACAGCCCGAG ATACACACGG GAAGATTATT AGTTGGTACT TTTGTGGCTC TTTGTTTCAA   480
                    CCTGTTCACG ATGCTTTCTA TTCGGAATAA ACCTTTTGCT TATGTCTCAG AAGCTGCTAG   540
     35             TACGGACTGG CTTCAAGAGC ATGTTGCAGA TCTTAGTCGA AGCTTGTGTG GAATTATTCC   600
                    GGGACTTAGC AGTATCTTCC TTCCCCGAAT GAATCCATTT GTTTTGATTG ATCTTGCTGG   660
                    AGCATTTGCT CTTTGTATTA CATATATGCT CATTGAAATT AATAATTATT TTGCCGTAGA   720
                    CACTGCCTCT GCTATAGCTA TTGCCTTGAT GACATTGGC ACTATGTATC CCATGAGTGT   780
                    GTACAGTGGG AAAGTCTTAC TCCAGACAAC ACCACCCCAT GTTATTGGTC AGTTGGACAA   840
     40             ACTCATCAGA GAGGTATCTA CCTTAGATGG AGTTTTAGAA GTCCGAAATG AACATTTTTG   900
                    GACCCTAGGT TTTGGCTCAT TGGCTGGATC AGTGCATGTA AGAATTCGAC GAGATGCCAA   960
                    TGAACAAATG GTTCTTGCTC ATGTGACCAA CAGGCTGTAC ACTCTAGTGT CTACTCTAAC  1020
                    TGTTCAAATT TTCAAGGATG ACTGGATTAG GCCTGCCTTA TTGTCTGGGC CTGTTGCAGC  1080
                    CAATGTCCTA AACTTTTCAG ATCATCACGT AATCCCAATG CCTCTTTTAA AGGGTACTGA  1140
     45             TGATTTGAAC CCAGTTACAT CAACTCCAGC TAAACCTAGT AGTCCACCTC CAGAATTTTC  1200
                    ATTTAACACT CCTGGGAAAA ATGTGAACCC AGTTATTCTT CTAAACACAC AAACAAGGCC  1260
                    TTATGGTTTT GGTCTCAATC ATGGACACAC ACCTTACAGC AGCATGCTTA ATCAAGGACT  1320
                    TGGAGTTCCA GGAATTGGAG CAACTCAAGG ATTGAGGACT GGTTTTACAA ATATACCAAG  1380
                    TAGATATGGA ACTAATAATA GAATTGGACA ACCAAGACCA TGATAGACTC TAACTTATTT  1440
     50             TTATAAGGAA TATTGACTCC TTGGCTTCCA ATTTATTTAG TAATCCAACT TTGCATTGAC  1500
                    TGTTTAATCA TTTACTCTAA ATGTTAGATA ATAGTAGTCT TGTTCACATT TCATGAAACC  1560
                    TATGAAACTA TATTTTGTA AAAATGTATT GTGACAGTGA AATCCTCGTA AATGTTAAAG  1620
                    GCTTTAAATA GGCTTCCTTT AGAAAATGTG TTTCTTTAAA TTTGGATTTT GGTATCTTTG  1680
                    GTTTTGTAGT TGACTGCAGT GTGATGTGAC CTTACCTTTA TAAGAGCCAC TTGATGGAGT  1740
     55             AGATCTGTCA CATTACTAAG ATACGATATT TCTTTTTTTT TCCGAGACGG AGTCTTGCTC  1800
                    TGCCACTGTG CCCGGCCAAT ACATTATTAT TAACTTAAGG CTGTACTTTA TTAAGGCTTC  1860
                    CTTAGTTTTT GTTTTGTTTT GTTTTTTGAG ATGGAGTCTC ACTCTGTCGC CCAGGCTGGA  1920
                    ATGCAGTGGC ATGATCTCAG CTCACTGCAA CCTCTGCCTC CTGAGTTCAA ATGATTCTCC  1980
                    TGCCTCAGCC TCCCGAGTAG CTGGGATTAC AGGCACCTGC CACCACGCCC AGCTAATTTT  2040
     60             TGTATTTTTA GTAAAGACGG GGGATTTCAC CATGTTGGCC AGGCTGGTCT TGAACTCCTG  2100
                    ACCTCATGAT CCACCCACCT TAGCCTCCCA AAGTGCTGGG ATTAGGTGTG AGCCACCGCA  2160
                    CCTGGCCGAT ATTTTCTTTA ATGAAATTTA TAAATATGCT TCTTGAATAA TACACATTTT  2220
                    GGGAAAGGGA AAAATGTCTG TTCAAAAAGT AAAGGTCTCT TTTATAGCTT TTCCAAACTT  2280
                    AATTGCTAAA TTTTTCTTTG AGGTTCTCCT GAATTATGTC TTACAAACTA AAAGCAAAAA  2340
     65             TTTTTAGCAG AAATTTTGGA ATACATTCTA TCTAGCACAA TTTGAATTTT TAATTATCAA  2400
                    GATTTTTGTT AAAGTTTCTC TCCTTTAAAA ATTTTAGTAC ATTTGTAAAT

                    Seq ID NO: 205 Protein sequence
                    Protein Accession #: BAB70980.1


     70             1          11         21         31         41         51
                    |          |          |          |          |          |
                    MGTIHLFRKP QRSFPGKLLR EPRLVAADRR SWKILLFGVI NLICTGFLLM WCSSTNSIAL    60
                    TAYTYLTIED LPSLMTCLIS YWVTLRKPSP VYSFGFERLE VLAVFASTVL AQLGALFILK   120
                    ESAERFLEQP EIHTGRLLVG TFVALCPNLF TMLSIRNKPF AYVSEAASTS WLQEHVADLS   180
     75             RSLCGIIPGL SSIFLPRMNP FVLIDLAGAF ALCITYMLIE INNYFAVDTA SAIAIALMTP   240
                    GTMYPMSVYS GKVLLQTTPP RVIGQLDKLI REVSTLDGVL EVRNEHFWTL GFGSLAGSVH   300
                    VRIRRDANEQ MVLAHVINRL YTLVSTLTVQ IFKDDWIRPA LLSGPVAANV LNFSDHRVIP   360
                    MPLLKGTDDL NPVTSTPAKP SSPPPEFSFN TPGKNVNPVI LLNTQTRPYG FGLNHGHTPY   420
                    SSMLNQGLGV PGIGATQGLR TGFTNIPSRY GINNRIGQPR P

     80             Seq ID NO: 206 DNA sequence
                    Nucleic Acid Accession #: NM_016361.1
                    Coding sequence: 397..1662

                    1          11         21         31         41         51
```

```
        |         |         |         |         |         |
        GGAACTCAGG GCCGGCTCCT GTTCCTTCAA GAGTGCTGGA GGCCAAACTT GAAATACAAG    60
        TTTAATGTTC CTCGTCGGGC AAAAGATAAG GATCCGATCT CCCCCGGCCC GGTGTGCAGC   120
        AGGAGCGACC AACCCCGACC CGGGTTAAAA CTCCCAGGGA CTCTTCGCTG CTGCCACCTC   180
        TTGTTCTCTC CCCCGTTCCC ACTCGGGGTC TCCCTCAGGG CCGGGAGGCA CAGCGGTCCC   240
        TGCTTGCTGA AGGGCTGGAT GTACGCATCC GCAGGTTCCC GCGGACTTGG GGGCGGCCGC   300
        TGAGCCCCGG CGCCCGCAGA AGACTTGTGT TTGCCTCCTG CAGCCTCAAC CCGGAGGCAG   360
        CGAGGGCCTA CCACCATGAT CACTGGTGTG TTCAGCATGC GCTTGTGGAC CCCAGTGGGC   420
        GTCCTGACCT CGCTGGCGTA CTGCCTGCAC CAGCGGCGGG TGGCCCTGGC CGAGCTGCAG   480
        GAGGCCGATG GCCAGTGTCC GGTCGACCGC AGCCTGCTGA AGTTGAAAAT GGTGCAGGTC   540
        GTGTTTCGAC ACGGGGCTCG GAGTCCTCTC AAGCCGCTCC CGCTGGAGGA GCAGGTAGAG   600
        TGGAACCCCC AGCTATTAGA GGTCCCACCC CAAACTCAGT TTGATTACAC AGTCACCAAT   660
        CTAGCTGGTG GTCCGAAACC ATATTCTCCT TACGACTCTC AATACCATGA GACCACCCTG   720
        AAGGGGGGCA TGTTTGCTGG GCAGCTGACC AAGGTGGGCA TGCAGCAAAT GTTTGCCTTG   780
        GGAGAGAGAC TGAGGAAGAA CTATGTGGAA GACATTCCCT TTCTTTCACC AACCTTCAAC   840
        CCACAGGAGG TCTTTATTCG TTCCACTAAC ATTTTTCGGA ATCTGGAGTC CACCCGTTGT   900
        TTGCTGGCTG GGCTTTTCCA GTGTCAGAAA GAAGGACCCA TCATCATCCA CACTGATGAA   960
        GCAGATTCAG AAGTCTTGTA TCCCAACTAC CAAAGCTGCT GGAGCCTGAG GCAGAGAACC  1020
        AGAGGCCGGA GGCAGACTGC CTCTTTACAG CCAGGAATCT CAGAGGATTT GAAAAAGGTG  1080
        AAGGACAGGA TGGGCATTGA CAGTAGTGAT AAAGTGGACT TCTTCATCCT CCTGGACAAC  1140
        GTGGCTGCCG AGCAGGCACA CAACCTCCCA AGCTGCCCCA TGCTGAAGAG ATTTGCACGG  1200
        ATGATCGAAC AGAGAGCTGT GGACACATCC TTGTACATAC TGCCCAAGGA AGACAGGGAA  1260
        AGTCTTCAGA TGGCAGTAGG CCCATTCCTC CACATCCTAG AGAGCAACCT GCTGAAAGCC  1320
        ATGGACTCTG CCACTGCCCC CGACAAGATC AGAAAGCTGT ATCTCTATGC GGCTCATGAT  1380
        GTGACCTTCA TACCGCTCTT AATGACCCTG GGGATTTTTG ACCACAAATG GCCACCGTTT  1440
        GCTGTTGACC TGACCATGGA ACTTTACCAG CACCTGGAAT CTAAGGAGTG GTTTGTGCAG  1500
        CTCTATTACC ACGGGAAGGA GCAGGTGCCG AGAGGTTGCC CTGATGGCCT CTGCCCGCTG  1560
        GACATGTTCT TGAATGCCAT GTCAGTTTAT ACCTTAAGCC CAGAAAAATA CCATGCACTC  1620
        TGCTCTCAAA CTCAGGTGAT GGAAGTTGGA AATGAAGAGT AACTGATTTA TAAAAGCAGG  1680
        ATGTGTTGAT TTTAAAATAA AGTGCCTTTA TACAAAAAAA AAAAAAAAAA A
```

Seq ID NO: 207 <u>Protein sequence</u>
Protein Accession #: NP_057445.1

```
        1          11        21        31        41        51
        |          |         |         |         |         |
        MRLWTPVGVL TSLAYCLHQR RVALAELQEA DGQCPVDRSL LKLKMVQVVF RHGARSPLKP    60
        LPLEEQVEWN PQLLEVPPQT QFDYTVTNLA GGPKPYSPYD SQYHETTLKG GMFAGQLTKV   120
        GMQQMFALGE RLRKNYVEDI PFLSPTFNPQ EVFIRSTNIF RNLESTRCLL AGLFQCQKEG   180
        PIIIHTDEAD SEVLYPNYQS CWSLRQRTRG RRQTASLQPG ISEDLKKVKD RMGIDSSDKV   240
        DFFILLDNVA AEQAHNLPSC PMLKRFARMI EQRAVDTSLY ILPKEDRESL QMAVGPFLHI   300
        LESNLLKAMD SATAPDKIRK LYLYAAHDVT FIPLLMTLGI FDHKWPPFAV DLTMELYQHL   360
        ESKEWFVQLY YHGKEQVPRG CPDGLCPLDM FLNAMSVYTL SPEKYHALCS QTQVMEVGNE   420
        E
```

Seq ID NO: 208 <u>DNA sequence</u>
Nucleic Acid Accession #: CAT cluster

```
        1          11        21        31        41        51
        |          |         |         |         |         |
        TTTGAGGGGG TGGTGGGGCG AGTTTAATTC ATAAAGAAGC CTCCTGATCA GAAAGGGGCC    60
        TAACAGCCTG CCCTTGGAGA GAAGTCCTTC CTTGAGGATA AGGCCTCCCA GGGGAGGAGG   120
        TGCTGGGGGC CAGTGTTAGG CTTCAGGCCA TCCCTGGAGG CCAGTCCTGT GCTCAGCAAG   180
        TAGTGGCAGA GCCTGGAGTG ATGAGTGGGA TGGCCTTCTC AGGTACAGGA CTGTGCTGCT   240
        TCTGGCTGCT CTTGCATTTG CATTTGCCAC TCAGAACTGC CGCGGATCCCA GCAATGGCCA   300
        GGAGCCCTCC GCAGATCAGT CGGCTCAGCT GCAGGTTTTT CCAGTCATAG TAGAAGGGAT   360
        CGTCTTTATT GGCAAATGGG TCATTGGCTT CCAAGGCAGT CAGGCCAACT GTGTGACTCT   420
        GCAGGTTCCT CACTGCTCCT TCACCAGTGT CCTGCGAGGT CACCTTGGCG AGGGCTCACC   480
        TGAGCTGGCA GCGCAG
```

Seq ID NO: 209 <u>DNA sequence</u>
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..564

```
        1          11        21        31        41        51
        |          |         |         |         |         |
        ATGGAGCCCT GGGCGTGGCT GCAGGGTTTA AAGAGCCGAC CCACGTGCCC AGCAGCCTCC    60
        TCAGATCCGT TCTCTGCGCT GCCAGCTCAG GACACTGGTG AAGGAGCAGT GAGGAACCTG   120
        CAGAGTCACA CAGTTGGCCT GACTGCCTTG GAAGCCAATG ACCCATTTGC CAATAAAGAC   180
        GATCCCTTCT ACTATGACTG GAAAAACCTG CAGCTGAGCG GACTGATCTG CGGAGGGCTC   240
        CTGGCCATTG CTGGGATCGC GGCAGTTCTG AGTGGCAAAT GCAAATGCAA GAGCAGCCAG   300
        AAGCAGCACA GTCCTGTACC TGAGAAGGCC ATCCCAATTCA CAGATTCTC  360
        ACCTTGGCCA AATCAAATAA ACCTTTATCT CCAAGCACCT TTGTCTTGGT GTTTGGCATC   420
        AGCTACACAT CAGTCTTCCG AGTGCCTCTT TCTGCGTCCC TGTACCCTGC CATTCCTGGT   480
        GATGCTGCTG CCCTCACATC AGGCCATCCA AGCATGCAGA ACATAAGCAT GCAGAACACT   540
        GGAACGAAGG GCTGTACCTA ATGA
```

Seq ID NO: 210 <u>Protein sequence</u>
Protein Accession #: FGENESH predicted

```
        1          11        21        31        41        51
        |          |         |         |         |         |
        MEPWAWLQGL KSRPTCPAAS SDPFSALPAQ DTGEGAVRNL QSHTVGLTAL EANDPFANKD    60
        DPFYYDWKNL QLSGLICGGL LAIAGIAAVL SGKCKCKSSQ KQHSPVPEKA IPLITPGRFL   120
```

```
TLAKSNKPLS PSTPVLVFGI SYTSVFRVPL SASLYPAIPG DAAALTSGHP SMQNISMQNT   180
GTKGCT
```

Seq ID NO: 211 <u>DNA sequence</u>
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..318

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCCCGGCC ACCCCGTCTG TGAAGTGAGG AGCACCTCTG CCCGGCTGCC CCGTCTGGGA   60
AGTGAGGAGC GCCTCTGCCC GGCTGCCACC CCGTCTGTGA GTGCCTGCTG CGCTGGGCCC   120
AGGCCGCCCG TGCCCTGCCA GGCCCTCCGG CCCCCCACCT TCCACCCCAG GGCCTGCTCC   180
TCACCCCAGG GTCCATCTC CCTAGTTTCC ACCAGAGACT GGGTCTTCAT TCTCACCCTG   240
CTACACAGCC CCTACCAGAA CGTTCTGAAA TGCAAACCTA ACAACTGTCT CACCCCAGCA   300
GGAAACTCCC CAGGGTCCCG GGCCCCCTGC GGGGTTGCAG GCCTCACTCT TCGCGCCCAT   360
CCCTCCGCCC TGACCGCCCT GAGCTCGCCC CCAGTGCTGG CCCTTCACGT CCAGTTATCC   420
CTCCCAGCCT CCAAGGTCCC CGTTACCGAA GACCGCCACC ATCACGACAT AGCGCAGCAC   480
ATATGGGACA CTGGTGAAGG AGCAGTGAGG AACCTGCAGA GTCACACAGT TGGCCTGACT   540
GCCTTGGAAG CCAATGACCC ATTTGCCAAT AAAGACGATC CCTTCTACTA TGACTGGAAA   600
AACCTGCAGC TGAGCGGACT GATCTGCGGA GGGCTCCTGG CCATTGCTGG GATCGCGGCA   660
GTTCTGAGTG GCAAATGCAA ATGCAAGAGC AGCCAGAAGC AGCACAGTCC TGTACCTGAG   720
AAGGCCATCC CACTCATCAC TCCAGGCAGA TTTCTCACCT TGGCCAAATC AAATAAACCT   780
TTATCTCCAA GCACCTTTGT CTTGGTGTTT GGCATCAGCT ACACATCAGT CTTCCGAGTG   840
CCTCTTTCTG CGTCCCTGTA CCCTGCCATT CCTGGTGATG CTGCTGCCCT CACATCAGGC   900
CATCCAAGCA TGCAGAACAT AAGCATGCAG AACACTGGAA CGAAGGGCTG TACCTAA
```

Seq ID NO: 212 <u>Protein sequence</u>
Protein Accession #: FGENESH predicted

```
1          11         21         31         41         51
|          |          |          |          |          |
MPGHPVCEVR STSARLPRLG SEERLCPAAT PSVSACCAGP RPPVPCQALR PPTFHPRACS   60
SPQGSISLVS TRDWVFILTL LHSPYQNVLK CKPNNCLTPA GNSPGSRAPC GVAGLTLRAH   120
PSALTALSSP PVLALHVQLS LPASKVPVTE DRHHHDIAQH IWDTGEGAVR NLQSHTVGLT   180
ALEANDPFAN KDDPFYYDWK NLQLSGLICG GLLAIAGIAA VLSGKCKCKS SQKQHSPVPE   240
KAIPLITPGR FLTLAKSNKP LSPSTFVLVF GISYTSVFRV PLSASLYPAI PGDAAALTSG   300
HPSMQNISMQ NTGTKGCT
```

Seq ID NO: 213 <u>DNA sequence</u>
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..1758

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGATGGGGT CTCATGTTGC CCAGGCTGGT CTTGAACTCC TGGGCTCGAG TGACCCTCCT   60
GCCTTGGCCT CCGAAAGTGC TGGGATTACA GGACTGTTAT TACAGGAATC CATAACACTG   120
GAGGATGTGG CTGTGGACTT CACTTGGGAG GAGTGGCAAC TCCTGGGCGC TGCTCAGAAG   180
GACCTGTACC GGGATGTGAT GTTGGAGAAC TACAGCAACC TGGTGGCAGT GGGGTATCAA   240
GCCAGCAAAC CGGATGCACT CTTCAAGTTG GAACAAGGAG AACAACTGTG GACAATTGAA   300
GATGGAATCC ACAGTGGAGC CTGTTCAGGT TCTCCAAAGG TCCCGTTCTC CATTTTCTCA   360
TCTGTGCCCT TCACTCTTCA AAATTGCCTT CATTCTAACA TATGGAAAGT TGATCATGTG   420
CTGGAGCGCT TGCAGAGTGA AAGCCTGGTG AACAGAAGGA AACCATGTCA TGAACATGAT   480
GCATTTGAAA ATATTGTTCA TTGCAGCAAA AGTCAGTTTC TGTTAGGGCA AAATCATGAT   540
ATATTTGACT TACGTGGAAA AAGTTTGAAA TCCAATTTAA CTTTAGTTAA CCAGAGCAAA   600
GGCTATGAAA TAAAGAACTC TGTTGAGTTT ACTGGAAATG GGGACTCCTT TCTTCATGCT   660
AACCATGAAC GACTTCATAC TGCAATTAAA TTCCCTGCAA GTCAAAAACT CATCAGCACT   720
AAGTCCCAAT TCATCAGTCC CAAGCATCAG AAAACACGAA AATTAGAGAA GCATCATGTG   780
TGCAGTGAAT GTGGGAAAGC CTTCATCAAG AAGTCTTGGC TAACTGATCA CCAGGTAATG   840
CATACAGGAG AGAAACCCCA CAGATGTAGT CTATGTGAGA AAGCCTTCTC CAGAAAGTTC   900
ATGCTTACTG AACATCAGCG AACTCATACA GGAGAAAAAC CTTATGAATG CCCTGAATGT   960
GGCAAAGCCT TTCTCAAGAA ATCACGGCTC AACATACATC AGAAAACACA TACGGGAGAG   1020
AAACCCTATA TATGCAGTGA ATGTGAAAA GGCTTCATCC AGAAAGGAAA TCTCATTGTA   1080
CACCAGCGAA TTCATACAGG TGAGAAACCT TATATATGCA ATGAATGTGG AAAAGGCTTC   1140
ATTCAGAAGA CGTGTCTCAT AGCACATCAG AGATTTCACA CAGGAAAGAC GCCCTTTGTG   1200
TGCAGTGAAT GTGGAAAATC CTGTTCTCAG AAATCAGGTC TCATTAAACA TCAAAGAATT   1260
CACACAGGAG AGAAACCCTT TGAATGTAGT GAATGTGGGA AAGCCTTTAG CACAAAGCAA   1320
AAGCTCATTG TCCATCAAAG GACTCATACA GGAGAGAAAC CTATGGCTG TAACGAGTGT   1380
GGGAAAGCGT TTGCGTATAT GTCGTGTCTG GTTAAGCATA AGAGAATACA CACAAGGGAG   1440
AAACAAGAGG CAGCCAAGGT GGAAAATCCT CCTGCAGAGA GGCACAGCTC ATTACACACC   1500
AGTGATGTCA TGCAGGAGAA AAACTCTGCT AACGGGGCGA CTACACAAGT GCCTTCTGTG   1560
GCCCCTCAGA CATCATTAAA CATCAGCGGC CTCCTGGCAA ACAGGAACGT AGTCCTGTG   1620
GGACAGCCAG TGGTCAGATG TGCAGCCTCA GGAGATAACA GAGGATTGC ACAGGACAGA   1680
AACCTTGTGA ATGCAGTGAA TGTGGTTGTG CCTTCCGTGA TCAATTATGT CTTATTTTAT   1740
GTTACAGAAA ACCCATAG
```

Seq ID NO: 214 <u>Protein sequence</u>
Protein Accession #: FGENESH predicted

```
1          11         21         31         41         51
|          |          |          |          |          |
MMGSHVAQAG LELLGSSDPP ALASESAGIT GLLLQESITL EDVAVDFTWE EWQLLGAAQK   60
DLYRDVMLEN YSNLVAVGYQ ASKPDALFKL EQGEQLWTIE DGIHSGACEG SPKVPPSIFS   120
SVPFTLQNCL HSNIWKVDHV LERLQSESLV NRRKPCHEHD AFENIVHCSK SQFLLGQNHD   180
IFDLRGKSLK SNLTLVNQSK GYEIKNSVEF TGNGDSFLHA NHERLHTAIK FPASQKLIST   240
```

```
          KSQPISPKHQ KTRKLEKHHV CSECGKAFIK KSWLTDHQVM HTGEKPHRCS LCEKAFSRKF  300
          MLTEHQRTHT GEKPYECPEC GKAFLKKSRL NIHQRTHTGE KPYICSECGK GFIQKGNLIV  360
          HQRIHTGEKP YICNECGKGF IQKTCLIAHQ RFHTGKTPFV CSECGKSCSQ KSGLIKHQRI  420
          HTGEKPFECS ECGKAFSTKQ KLIVHQRTHT GERPYGCNEC GKAFAYMSCL VKHKRIHTRE  480
          KQEAAKVENP PAERHSSLHT SDVMQEKNSA NGATTQVPSV APQTSLNISG LLANRNVVLV  540
          GQPVVRCAAS GDNRGPAQDR NLVNAVNVVV PSVINYVLFY VTENP
```

Seq ID NO: 215 DNA sequence
Nucleic Acid Accession #: NM_032190.1
Coding sequence: 502..1332

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          GATTCCGTGT TCTTGGCCAT GTTAGCCATA ATATCCTGTG CAGTATGTTT TTCCTGTGCA    60
          GAGGCAAAAA CATATTGGGC ATATGTTCCC AAGCCCCCAG CAGTATGACC CATACTTTGG   120
          AGTGACACTC CTCCTAAGAT TTATCATGAT TAAGGAGCAT GGGCTCCAGG ACCCCTAACT   180
          CCACCTGACA TAGAACAGTT AGACTCTCAG AATAATGTCA TTAATTATAC CGCTCCATTG   240
          GAAGGACTTC CTTTGTGTGT CACCCACAAG ACATCACTCA GCCATAGCTG TCTTACAGTT   300
          CAAGCTCACA CATGGTTGAG TCACTATGGG AAAATCATGT ACTTATTAAG TCTTGGTTAT   360
          ATTAATGTAA CCGGTGTGCT AACCAACCAT TCCTGGCCCA ATCGCCTTCA TTGTGCTGAC   420
          TATACAGAAT GGATTCCCTT CAATAGTTCC TACCCCCCTC CATAGACCCA GTGTCTTGGC   480
          CCACTGGCTA GAAAACAATC TATGTTAACT GGAGACATTG TGGATTGGGG ACCTAAAGGC   540
          CAATTAGATG GAAAAGAAGA AAATCAGAAA TCGTGGCACA AACTTTGCTG GCATTGGTGG   600
          CAAGCTTTTA ATGCTTCTTC TTTATATAAC ACTGGGATCC AATCCCAGTC GGCCGCCCAG   660
          ATTGCTTGGC ATGGAGCAGG CTTTAGCCCG CCTCTTCCTA AGTGGCATTA TCTAGGGAGG   720
          AAAGGACCAA TTCAAAAGAT GATATGGAAG GCAGCATTCC CATTTATGAA TGGCAACATC   780
          TGGGTTGCCA TAATACTATC CAATAATAGC AATAGTAAGC AACACAGTCT TAATGTTACA   840
          TTTGTAAAGA ATATCACCAC TCAATTTACA GTTTGTGTTT TTAATCCTTA TGTGTTTTTG   900
          GCAGCTAAGA AGGACCAGCT CCAGGTAAAC AATACCCAAT TGACCTGTAA ATCTTGCCAG   960
          TTATATCACT GCATTAATCA TAGCACATTG CAAACACATA ATATCTCTAC TTTGATGATT  1020
          TTAGGTTGCA TCCCTGGGCT ATGGATTCCT GTTAATCTGT CTGAGCCATG GGCTGCCACA  1080
          ATTGCTTTAC ATTTTGTGAA ACTTCTTCTA ACTCAGTTTA CTCATTGTGT CCGTAGAGGC  1140
          TTAGGCATGA TAATTTTTGC TATTGTTTAC TTGGTCACAC TAATAATTTC TGTTGTGATG  1200
          TCCTCTGTAG CTTTGCATAG TTCTATTCAA ACAGCTCAGT ATGTGGAGAA CTGGACACGC  1260
          ACAGTCAACC AAGGGTGGCT ACTTGAGAAT AAAATTAACA CTGAGTTACA AACTGAAGTG  1320
          GCAGTGTTAT AATCCACGAT TCTATGGTTA GGGGAACAAG TACAAAGCTT GCAATTGCAG  1380
          CAGTAATTGT GTTGTCATTT TAATCACACT CATATTTGTG TAACCAACTT AGAATATAAC  1440
          CAAAGTGAGT ATCCATGGGA TCTTGTGAAA GCCCATTTGC AGGGAGCTTT CACATCCGAC  1500
          ATCACCTTTG ATATTGGTGA ATTACAAAAC AAAATTCTTG ATTTAAATAA ACAAATTCCA  1560
          GAGTTTCAGC CTTCTTTAGA AGACTGGACT GAATTCCCAGC AAGGCCTGAA GAGCGTCAAC  1620
          CCTTGGACCT ATCTAAAGCA CCACATTAAC ATCTTATATA TAGTTCTTGG AATAATGTTG  1680
          TTTTGTCTCT GTCTTCTGTT CATAGTCTGT AAAATCGGAT GGACTGCCAA TCGGAGAATG  1740
          AAAGCTACCC AGCCTGGCCT TACATTCTTT CACTTAATAC ATAAACAAGA AGGGGGAAAT  1800
          GTTGGGAGCC AAAAAGGCCA AAGGGATGGT GACCAACTCA GCATTCCACT GGAGGCTACA  1860
          TGATCAAACA GCAAACTGTT TATCATGAAT ACAGAATGTG GGCAAACTCG CTTCTGTGCC  1920
          TGCCCAGAAG GTTTGCTGAG GGCCATCGCT CCCTGGCCCC GGCTCCTTGA GGTTATCTGAC  1980
          TGGGACATCT AGAGCCTATT GTTCGAGGAA TGCAGTCTTG CAAGCCTACT CTGGACCGAG  2040
          CAGCTGACCT CTTCTTCCAC ACCCCTTCTC ACTATCTCTT TTGCCTAATA AATATGGAGG  2100
          GCTGTGTAAA GCTCAGGGCC CTTGTCCACT AGAGGCAAGG TGTCCCCTGA CCCTTCTTCC  2160
          AAACAT
```

Seq ID NO: 216 Protein sequence
Protein Accession #: NP_115566.1

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          MLTGDIVDWG PKGQLDGKEE NQKSWHKLCW HWWQAFNASS LYNTGIQSQS AAQIAWHGAG    60
          FSPPLPQWHY LGRKGPIQKM IWKAAFPFMN GNIWVAIILS MNSNSKQHSL NVTFVKNITT   120
          QFTVCVFNPY VFLAAKKDQL QVNNTQLTCK SCQLYHCINH STLQTHNIST LMILGCIPGL   180
          WIPVNLSEPW AATIALHFVK LLLTQFTHCV RRGLGMIIFA IVYLVTLIIS VVMSSVALHS   240
          SIQTAQYVEN WTRTVNQGWL LENKINTELQ TEVAVL
```

Seq ID NO: 217 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..1566

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          ATGGTGAACC CCAAATCCAC TTCCTCCCTC TTCAGGTTAT GTTTTTTGCT CCTGAGGAGT    60
          CAGAACCTGT GGGTTGAAGA GCAAATTCAA TGCAAAAACA TATTGGGCAT ATGTTCCCAA   120
          TCCCCCAGCA GTATGGCCTA TACTTTGGAG CTCACTCCTC CTGAGATTTA TCACGATCAG   180
          GGAGAGTGGG CTCCAGGACC CCTAACTCCC CGTGACATAG AAAAGTTAGA CTCTCAGAAC   240
          AATGTCATTA ATTATACCAC TCCACTGGAA GGACTCCCTT TGTTTATCAC CACAAAGACG   300
          TCGCTCAGCC ATAGCTGTCT TGCAATTCAA GCTCAAACAT GGTTGAGTCA CTATGGAAAA   360
          ATTATGTACT TATTAGGTCT TGGTTCTATT AATGTAACTG GTGTGCTAAC CAATCATTCC   420
          CAGTCCAGTC ACCCTAATTG TGCTGATTAT ACAGAATGGA TTCCATTCAA TAGTTCCTAC   480
          CCCACTCTGT GGACCCAGTG TCTTGATCCA CTGGCTAGTA AACAATATAT GTCAACTGAA   540
          GACACTGTGG ATTGGGAACC TAAAGGTCAA TTAGATGGAA AAGGTGAAAG TCAGAAATCA   600
          TGGCACAAAC TTCACTGGCA TTGGCGGCAA GCTTTTAATG CTTCTTCTTT ATACAACAGC   660
          AGAATCCAAT CCCAGTCTGC TGCTCAGATT GCTTGGCATG GAGCAGGCTT TAGCCCACCT   720
          CTTCCTCAGT TGCATTATCT GGGGAGGAAA GGACCAATTC AAGAAACTAT ATGGAAGGCA   780
          GCACTCCCAT TTATGAATGG CAACATCTGG ATTGGAACAC TGTCTAATAA TAGCAATAGT   840
          AAGCAACACA GTCTTAATGT TGCATTTGTA AAGAATATCA CCACTCAGTT TACAGTTTGT   900
          GTTTTTAATC CTTATGCCTT TTTGGCAGCT AAGAAGAACC AGCTTCAGGT GGAGAACTGG   960
```

```
ACACGCACAG CTGACCAAGC GAGGCTACTT CAGAATAAAA TTAACACTGA GTTACAAACT 1020
GAAGTGGCAA TGTTGAAATC CATGGTTCTG TGGTTAGGAG AACAGGTACA AAGCTTGCAG 1080
TTGCAGCAGC AATTGCGTCA TCATTTTAAT CACATTCATA TTTGCGTAAC TAACTCAGAA 1140
TATAACCAAA GTGAGTATCC GTGGGACCTT GTGAAAGCCC ATTTGCAAGG AGCTTTCACA 1200
TCCAACATCA CCTTTGATAT TGGTGAATTA CAAAACAAAA TTATTGATTT AAATAGGCAA 1260
ACTCAAGAAT TTCAGCCTTC TTTAGAAGAC TGGACCGAAT TCCAGGAAGG CCTGGAGAGC 1320
CTCAACCCTT GGACCTATCT AAGGCACCAC ATTAACATCT TATATGTAGT TCTTGGAATA 1380
ATGTTGTTTT GTCTCTGTCT TCGGTTTATA GTCTGTAAAA TCGGATGGAC CACCAATTGG 1440
AGAATGAGAG CCTCCCAGCC CAGCCTTACA TTCTTTCAAT TAATACATAA ACAGAAAGGG 1500
GGATATGCAG GGAGCCAAAG GCCTGTGGGA CGTGACCAAC TCAGCATTCT GCTGGAGGCT 1560
ATATGA
```

Seq ID NO: 218 Protein sequence
Protein Accession #: FGENESH predicted

```
1          11         21         31         41         51
|          |          |          |          |          |
MVNPKSTSSL FRLCPLLLRS QNLWVBBQIQ CKNILGICSQ SPSSMAYTLB LTPPEIYHDQ   60
GEWAPGPLTP RDIEKLDSQN NVINYTTPLB GLPLFITTKT SLSHSCLAIQ AQTWLSHYGK  120
IMYLLGLGSI NVTGVLTNHS QSSHPNCADY TEWIPFNSSY PTLWTQCLDP LASKQYMSTE  180
DTVDWBPKGQ LDGKGESQKS WHKLHWHWRQ AFNASSLYNS RIQSQSAAQI AWHGAGFSPP  240
LPQLHYLGRK GPIQETIWKA ALPFMNGNIW IGTLSNNSNS KQHSLNVAFV KNITTQFTVC  300
VFNPYAFLAA KKNQLQVENW TRTADQARLL QNKINTELQT EVAMLKSMVL WLGEQVQSLQ  360
LQQQLREHFN HIHICVTNSS YNQSEYPWDL VKAHLQGAFT SNITPDIGEL QNKIIDLNRQ  420
TQEFQPSLED WTEFQEGLES LNPWTYLRHH INILYVVLGI MLPCLCLRFI VCKIGWTTNW  480
RMRASQPSLT FFQLIHKQKG GYAGSQRPVG RDQLSILLEA I
```

Seq ID NO: 219 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..900

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCCGCCGC GGGAGCTGAG CGAGGCCGAG CCGCCCCCGC TCCGGCCCC GACCCCTCCC   60
CCGCGGCGGC GTAGCGCGCC CCCAGAGCTG GGCATCAAGT GCGTGCTGGT GGGCGACGGC  120
GCCGTGGGCA AGAGCAGCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGCGCTAC  180
CGGCCCACTG CGCTGGACAC CTTCTCTGGT ACGTACGTTC AATCGCCCGT GGGGCCGCGT  240
GGCTGCGGCG GGGCTGTGCA CCGGGGAGCT GGGGCGGGCG TCTCGGCGGG AGGGCGCAGA  300
GGACCCCGGG GAGGAGACTG GAGCAGGCCC CGAGGTGGCG CTGGTGCGGC CCAGGACGCT  360
CTTCCTAACT CAGGCTCTCC CCGCCCGCC CCTGCAGTGC AAGTCCTGGT GGATGGAGCT  420
CCGGTGCGCA TTGAGCTCTG GGACACAGCG GGACAGGAGG ATTTTGACCG ACTTCGTTCC  480
CTTTGCTACC CGGATACCGA TGTCTTCCTG GCGTGCTTCA GCGTGGTGCA GCCCAGCTCC  540
TTTCAAAACA TCACAGAGAA ATGGCTGCCC GAGATCCGCA CGCACAACCC CCAGGCGCCT  600
GTGCTGCTGG TGGGCACCCA GGCCGACCTG AGGGACGATG TCAACGTACT AATTCAGCTG  660
GACCAGGGGG GCCGGGAGGG CCCCGTGCCC CAACCCCAGG CTCAGGGGTC GGCCGAGAAG  720
ATCCGAGCCT GCTGCTACCT TGAGTGCTCA GCCTTGACGC AGAAGAACTT GAAGGAAGTA  780
TTTGACTCGG CTATTCTCAG TGCCATTGAG CACAAAGCCC GGCTGGAGAA GAAACTGAAT  840
GCCAAAGGTG TGCGCACCCT CTCCCGCTGC CGCTGGAAGA AGTTCTTCTG CTTCGTTTGA
```

Seq ID NO: 220 Protein sequence
Protein Accession #: FGENESH predicted

```
1          11         21         31         41         51
|          |          |          |          |          |
MPFRELSEAE PPPLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY   60
RPTALDTFSG TYVQSPVRFR GCGGAVHRGA GAGVSAGGRR GPRGGDWSRP EGGAGAAQDA  120
LPNSGSPRPA PAVQVLVDGA PVRIELWDTA GQEDFDRLRS LCYPDTDVFL ACFSVVQPSS  180
FQNITEKWLP EIRTHNPQAP VLLVGTQADL RDDVNVLIQL DQGGREGFVP QPQAQGLAEK  240
IRACCYLECS ALTQKNLKEV FDSAILSAIE HKARLEKKLN AKGVRTLSRC RWKKFFCFV
```

Seq ID NO: 221 DNA sequence
Nucleic Acid Accession #: XM_063832.2
Coding sequence: 1..711

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCCGCCGC GGGAGCTGAG CGAGGCCGAG CCGCCCCCGC TCCGGGCCCC GACCCCTCCC   60
CCGCGGCGGC GTAGCGCGCC CCCAGAGCTG GGCATCAAGT GCGTGCTGGT GGGCGACGGC  120
GCCGTGGGCA AGAGCAGCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGCGCTAC  180
CGGCCCACTG CGCTGGACAC CTTCTCTGTG CAAGTCCTGG TGGATGGAGC TCCGGTGCGC  240
ATTGAGCTCT GGGACACAGC GGGACAGGAG GATTTTGACC GACTTCGTTC CCTTTGCTAC  300
CCGGATACCG ATGTCTTCCT GGCGTGCTTC AGCGTGGTGC CTTTCAAAAC  360
ATCACAGAGA AATGGCTGCC CGAGATCCGC ACGCACAACC CCCAGGCGCC TGTGCTGCTG  420
GTGGGCACCC AGGCCGACCT GAGGGACGAT GTCAACGTAC TAATTCAGCT GGACCAGGGG  480
GGCCGGGAGG GCCCCGTGCC CCAACCCCAG GCTCAGGGTC TGGCCGAGAA GATCCGAGCC  540
TGCTGCTACC TTGAGTGCTC AGCCTTGACG CAGAAGAACT TGAAGGAAGT ATTTGACTCG  600
GCTATTCTCA GTGCCATTGA GCACAAAGCC CGGCTGGAGA AGAAACTGAA TGCCAAAGGT  660
GTGCGCACCC TCTCCCGCTG CCGCTGGAAG AAGTTCTTCT GCTTCGTTTG A
```

Seq ID NO: 222 Protein sequence
Protein Accession #: XP_063832.1

```
1          11         21         31         41         51
|          |          |          |          |          |
```

```
MPPRELSEAB PPPLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY   60
RPTALDTFSV QVLVDGAPVR IELWDTAGQE DPDRLRSLCY PDTDVFLACF SVVQPSSFQN  120
ITBKWLPEIR TBNPQAPVLL VGTQADLRDD VNVLIQLDQG GRESJPVPQPQ AQGLAEKIRA  180
CCYLECSALT QKNLKEVFDS AILSAIEHKA RLEKKLNAKG VRTLSRCRWK KPPCFV
```

Seq ID NO: 223 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..1161

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG   60
TTCCGAGATG ACTTCATTGC CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTTATCTTT  120
GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA  180
TCCAGCCGGA TTTTCCTGTT CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG  240
CCGTTCGTGA TGGACTACTA TGTGCGGCGT TCAGACTGGA AGTTTGGGGA CATCCCTTGC  300
CGGCTGGTGC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG  360
GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC  420
AATTGGACAG CAGCCATCAT CTCTTGCCTT CTGTGGGGCA TCACTGTTGG CCTAACAGTC  480
CACCTCCTGA AGAAGAAGTT GCTGATCCAG AATGGCCCTG CAAATGTGTG CATCAGCTTC  540
AGCATCTGCC ATACCTTCCG GTGGCACGAA GCTATGTTCC TCCTGGAGTT CCTCCTGCCC  600
CTGGGCATCA TCCTGTTCTG CTCAGCCAGA ATTATCTGGA GCCTGCGGCA GAGACAAATG  660
GACCGGCATG CCAAGATCAA GAGAGCCATC ACCTTCATCA TGGTGGTGGC CATCGTCTTT  720
GTCATCTGCT TCCTTCCCAG CGTGGTTGTG CGGATCCGCA TCTTCTGGCT CCTGCACACT  780
TCGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCGTTCTT TATCACTCTC  840
AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTACTTCTC CAGCCCATCC  900
TTTCCCAACT TCTTCTCCAC TTTGATCAAC CGCTGCCTCC AGAGGAAGAT GACAGGTGAG  960
CCAGATAATA ACCGCAGCAC GAGCGTCGAG CTCACAGGGG ACCCCAACAA AACCAGAGGC 1020
GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCCTCTTA TCTGGGCCCA 1080
ACCTCAAATA ACCATTCCAA GAAGGGACAT TGTCACCAAG AACCAGCATC TCTGGAGAAA 1140
CAGTTGGGAT GTTGCATCGA G
```

Seq ID NO: 224 Protein sequence
Protein Accession #: Eos sequence

```
1          11         21         31         41         51
|          |          |          |          |          |
MNRHHLQDHF LEIDKKNCCV FRDDFIAKVL PPVLGLEFIF GLLGNGLALW IPCFHLKSWK   60
SSRIFLFNLA VADFLLIICL PFVMDYYVRR SDWKFGDIPC RLVLFMFAMN RQGSIIFLTV  120
VAVDRYFRVV HPHHALNKIS NWTAAIISCL LWGITVGLTV HLLKKKLLIQ NGPANVCISF  180
SICHTFRWHE AMFLLEFLLP LGIILFCSAR IIWSLRQRQM DRHAKIKRAI TFIMVVAIVF  240
VICFLPSVVV RIRIFWLLHT SGTQNCEVYR SVDLAFFITL SFTYMNSMLD PVVYYFSSPS  300
FPNFFSTLIN RCLQRKMTGE PDNNRSTSVE LTGDPNKTRG APEALMANSG EPWSPSYLGP  360
TSNNHSKKGH CHQEPASLEK QLGCCIE
```

Seq ID NO: 225 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..1092

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG   60
TTCCGAGATG ACTTCATTGT CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTTATCTTC  120
GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA  180
TCCAGCCGGA TTTTCCTGTT CAACCTGGCA GTGGCTGACT TTCTACTGAT CATCTGCCTG  240
CCCTTCCTGA TGGACAACTA TGTGAGGCGT TGGGACTGGA AGTTTGGGGA CATCCCTTGC  300
CGGCTGATGC TCTTCATGTT GGCTATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG  360
GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC  420
AATCGGACAG CAGCCATCAT CTCTTGCCTT CTGTGGGGCA TCACTATTGG CCTGACAGTC  480
CACCTCCTGA AGAAGAAGAT GCCGATCCAG AATGGCGGTG CAAATTGTG CAGCAGCTTC  540
AGCATCTGCC ATACCTTCCA GTGGCACGAA GCCATGTTCC TCCTGGAGTT CTTCCTGCCC  600
CTGGGCATCA TCCTGTTCTG CTCAGCCAGA ATTATCTGGA GCCTGCGGCA GAGACAAATG  660
GACCGGCATG CCAAGATCAA GAGAGCCATC ACCTTCATCA TGGTGGTGGC CATCGTCTTT  720
GTCATCTGCT TCCTTCCCAG CGTGGTTGTG CGGATCCGCA TCTTCTGGCT CCTGCACACT  780
TCGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCGTTCTT TATCACTCTC  840
AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTACTTCTC CAGCCCATCC  900
TTTCCCAACT TCTTCTCCAC TTTGATCAAC CGCTGCCTCC AGAGGAAGAT GACAGGTGAG  960
CCAGATAATA ACCGCAGCAC GAGCGTCGAG CTCACAGGGG ACCCCAACAA AACCAGAGGC 1020
GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCCTCTTA TCTGGGCCCA 1080
```

Seq ID NO: 226 Protein sequence
Protein Accession #: Eos sequence

```
1          11         21         31         41         51
|          |          |          |          |          |
MNRHHLQDHF LEIDKKNCCV FRDDFIVKVL PPVLGLEFIF GLLGNGLALW IPCFHLKSWK   60
SSRIFLFNLA VADFLLIICL PFLMDNYVRR WDWKFGDIPC RLMLFMLAMN RQGSIIFLTV  120
VAVDRYFRVV HPHHALNKIS NRTAAIISCL LWGITIGLTV HLLKKKMPIQ NGGANLCSSF  180
SICHTFQWHE AMFLLEFFLP LGIILFCSAR IIWSLRQRQM DRHAKIKRAI TFIMVVAIVF  240
VICFLPSVVV RIRIFWLLHT SGTQNCEVYR SVDLAFFITL SFTYMNSMLD PVVYYFSSPS  300
FPNFFSTLIN RCLQRKMTGE PDNNRSTSVE LTGDPNKTRG APEALMANSG EPWSPSYLGP  360
TSP
```

Seq ID NO: 227 DNA sequence
Nucleic Acid Accession #: NM_006018
Coding sequence: 61..1224

```
1          11         21         31         41         51
|          |          |          |          |          |
CGCCACTTTG CTGGAGCATT CACTAGGCGA GGCGCTCCAT CGGACTCACT AGCCGCACTC    60
ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG   120
TTCCGAGATG ACTTCATTGC CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTTATCTTT   180
GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA   240
TCCAGCCGGA TTTTCCTGTT CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG   300
CCGTTCGTGA TGGACTACTA TGTGCGGCGT TCAGACTGGA ACTTTGGGGA CATCCCTTGC   360
CGGCTGGTGC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG   420
GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC   480
AATTGGACAG CAGCCATCAT CTCTTGCCTT CTGTGGGGCA TCACTGTTGG CCTAACAGTC   540
CACCTCCTGA AGAAGAAGTT GCTGATCCAA AATGGCCCTG CAAATGTGTG CATCAGCTTC   600
AGCATCTGCC ATACCTTCCG GTGGCACGAA GCTATGTTCC TCCTGGAGTT CCTCCTGCCC   660
CTGGGCATCA TCCTGTTCTG CTCAGCCAGA ATTATCTGGA GCCTGCGGCA GAGACAAATG   720
GACCGGCATG CCAAGATCAA GAGAGCCATC ACCTTCATCA TGGTGGTGGC CATCGTCTTT   780
GTCATCTGCT TCCTTCCCAG CGTGGTTGTG CGGATCCGCA TCTTCTGGCT CCTGCACACT   840
TCGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCGTTCTT TATCACTCTC   900
AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTACTTCTC CAGCCCATCC   960
TTTCCCAACT TCTTCTCCAC TTTGATCAAC CGCTGCCTCC AGAGGAAGAT GACAGGTGAG  1020
CCAGATAATA ACCGGCAGCAC GAGCGTCGAG CTCACAGGGG ACCCCAACAA AACCGAGGC   1080
GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCCTCTTA TCTGGGCCCA  1140
ACCTCAAATA ACCATTCCAA GAAGGGACAT TGTCACCAAG AACCAGCATC TCTGGAGAAA  1200
CAGTTGGGCT GTTGCATCGA GTAATGTCAC TGGACTCGGC CTAAGGTTTC CTGGAACTTC  1260
CAGATTCAGA GAATCTGATT TAGGGAAACT GTGGCAGATG AGTGGGAGAC TGGTTGCAAG  1320
GTGTGACCAC AGGAATCCTG GAGGAACAGA GAGTAAAGCT TCTAGGCATC TGAAACTTGC  1380
TTCATCTCTG ACGCTCGCAG GACTGAAGAT GGGCAAATTG TAGGCGTTTC TGCTGAGCAG  1440
AGTTGGAGCC AGAGATCTAC TTGTGACTTG TTGGCCTTCT TCCCACATCT GCCTCAGACT  1500
GGGGGGGGCT CAGCTCCTCG GGTGATATCT AGCCTGCTTG TGAGCTCTAG CAGGGATAAG  1560
GAGAGCTGAG ATTGGAGGGA ATTGTGTTGC TCCTGGAGGA AGCCCAGGCA TCATTAAACA  1620
AGCCAGTAGG TCACCTGGCT TCCGTGGACC AATTCATCTT TCAGACAAGC TTTAGAGAAA  1680
TGGACTCAGG GAAGAGACTC ACATGCTTTG GTTAGTATCT GTGTTCCGG TGGGTGTAAT   1740
AGGGGATTAG CCCCAGAAGG GACTGAGCTA AACAGTGTTA TTATGGGAAA GGAAATGGCA  1800
TTGCTGCTTT CAACCAGCGA CTAATGCAAT CCATTCCTAT CTTGTTTATA GTAATCTAAG  1860
GGTTGAGCAG TTAAAACGGC TTCAGGATAG AAAGCTGTTT CCCACCTGTT TCGTTTTACC  1920
ATTAAAAGGG AAAGTGCCT CTGCCCCACG GGTAGAGGGG GTGCACGTTC CTCCTGGTTC   1980
CTTCGGCTTGT GTTTCTGTAC TTACCAAAAA TCTACCACTT CAATAAATTT TGATAGGAGA  2040
CAAAAAAAAA A
```

Seq ID NO: 228 Protein sequence
Protein Accession #: NP_006009.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MNRHHLQDHF LEIDKKNCCV FRDDFIAKVL PPVLGLEFIF GLLGNGLALW IPCFHLKSWK    60
SSRIFLFNLA VADFLLIICL PFVMDYYVRR SDWNFGDIPC RLVLFMFAMN RQGSIIFLTV   120
VAVDRYFRVV HPHHALNKIS NWTAAIISCL LWGITVGLTV HLLKKKLLIQ NGPANVCISF   180
SICHTFRWHE AMFLLEFLLP LGIILFCSAR IIWSLRQRQM DRHAKIKRAI TFIMVVAIVF   240
VICFLPSVVV RIRIFWLLHT SGTQNCEVYR SVDLAFFITL SFTYMNSMLD PVVYYFSSPS   300
FPNFFSTLIN RCLQRKMTGE PDNNRSTSVE LTGDPNKTRG APEALMANSG EPWSPSYLGP   360
TSNNHSKKGH CHQRPASLEK QLGCCIE
```

Seq ID NO: 229 DNA sequence
Nucleic Acid Accession #: NM_014398.1
Coding sequence: 64..1314

```
1          11         21         31         41         51
|          |          |          |          |          |
GGCACCGATT CGGGGCCTGC CCGGACTTCG CCGCACGCTG CAGAACCTCG CCCAGCCGCCC    60
ACCATCCCC GGCAGCTCAG CGCGGCGGCC GCGCTCTTCG CGTCCCTGGC CGTAATTTTG   120
CACGATGGCA GTCAAATGAG AGCAAAAGCA TTTCCAGAAA CCAGAGATTA TTCTCAACCT   180
ACTGCAGCAG CAACAGTACA GGACATAAAA AAACCTGTCC AGCAACCAGC TAAGCAAGCA   240
CCTCACCAAA CTTTAGCAGC AAGATTCATG GATGGTCATA TCACCTTTCA AACAGCGGAC   300
ACAGTAAAAA TTCCAACAAC TACCCCAGCA ACTACAAAAA ACACTGCAAC CACCAGCCCA   360
ATTACCTACA CCCTGGTCAC AACCCAGGCC ACACCCAACA ACTCACACAC AGCTCCTCCA   420
GTTACTGAAG TTACAGTCGG CCCTAGCTTA GCCCCTTATT GCCCACCACC CACCATCACC   480
CCACCAGCTC ATACAGCTGG AACCAGTTCA TCAACCGTCA GCCACACACC TGGGAACACC   540
ACTCAACCCA GTAACCAGAC CACCCTTCCA GCAACTTTAT CGATAGCACT GCACAAAAGC   600
ACAACCGGTC AGAAGCCTGA TCAACCCACC CATGCCCCAG GAACAAACGGC AGCTGCCCAC   660
AATACCACCC GCACAGCCTGC ACCTGCCTCC ACGGTTCCTG GGCCCACCCT TGCACCTCAG   720
CCATCGTCAG TCAAGACTGG AATTTATCAG GTTCTAAACG GAAGCAGACT CTGTATAAAA   780
GCAGAGATGG GGATACAGCT GATTGTTCAA GACAAGGAGT CGGTTTTTTC ACCTCGGAGA   840
TACTTCAACA TCGACCCCAA CGCAACGCAA GCCTCTGGGA ACTGTGGCAC CGGAAAATCC   900
AACCTTCTGT TGAATTTTCA GGGCGGGATTT GTGAATCTCA CATTTACCAA GGATGAAGAA   960
TCATATTATA TCAGTGAAGT GGGAGCCTAT TTGACCGTCT CAGATCCAGA GACAGTTTAC  1020
CAAGGAATCA AACATGCGGT GGTGATGTTC CAGACAGCAG TCGGGCATTC CTTCAAGTGC  1080
GTGAGTGAAC AGAGCCTCCA GTTGTCAGCC CACCTGCAGG TGAAACAAC CGATGTCCAA   1140
CTTCAAGCCT TTGATTTTGA AGATCACCAC TTTGGAAATG TGGATGAGTG CTCGTCTGAC  1200
TACACAATTG TGCTTCCTGT GATTGGGGCC ATCGTGGTTG GTCTCTGCCT TATGGGTATG  1260
GGTGTCTATA AAATCCGCCT AAGGTGTCAA TCATCTGGAT ACCAGAGAAT CTAATTGTTG  1320
CCCGGGGGGA ATGAAAATAA TGGAATTTAG AGAACTCTTT CATCCCTTCC AGGATGGGATG  1380
```

```
TTGGGAAATT  CCCTCAGAGT  GTGGGTCCTT  CAAACAATGT  AAACCACCAT  CTTCTATTCA  1440
AATGAAGTGA  GTCATGTGTG  ATTTAAGTTC  AGGCAGCACA  TCAATTTCTA  AATACTTTTT  1500
GTTTATTTTA  TGAAAGATAT  AGTGAGCTGT  TTATTTTCTA  GTTTCCTTTA  GAATATTTTA  1560
GCCACTCAAA  GTCAACATTT  GAGATATGTT  GAATTAACAT  AATATATGTA  AAGTAGAATA  1620
AGCCTTCAAA  TTATAAACCA  AGGGTCAATT  GTAACTAATA  CTACTGTGTG  TGCATTGAAG  1680
ATTTTATTTT  ACCCTTGATC  TTAACAAAGC  CTTTGCTTTG  TTATCAAATG  GACTTTCAGT  1740
GCTTTTACTA  TCTGTGTTTT  ATGGTTTCAT  GTAACATACA  TATTCCTGGT  GTAGCACTTA  1800
ACTCCTTTTC  CACTTTAAAT  TTGTTTTTGT  TTTTTGAGAC  GGAGTTTCAC  TCTTGTCACC  1860
CAGGCTGGAG  TACAGTGGCA  CGATCTCGGC  TTATGGCAAC  CTCCGCCTCC  CGGGTTCAAG  1920
TGATTCTCCT  GCTTCAGCTT  CCCGAGTAGC  TGGGATTACA  GGCACACACT  ACCACGCCTG  1980
GCTAATTTTT  GTATTTTTAT  TATAGACGGG  TTTCACCATG  TTGGCCAGAC  TGGTCTTGAA  2040
CTCTTGACCT  CAGGTGATCC  ACCCACCTCA  GCCTCCCAAA  GTGCTGGGAT  TACAGGCATG  2100
AGCCATTGCG  CCCGGCCTTA  AATGTTTTTT  TTAATCATCA  AAAAGAACAA  CATATCTCAG  2160
GTTGTCTAAG  TGTTTTTATG  TAAAACCAAC  AAAAAGAACA  AATCAGCTTA  TATTTTTTAT  2220
CTTGATGACT  CCTGCTCCAG  AATTGCTAGA  CTAAGAATTA  GGTGGCTACA  GATGGTAGAA  2280
CTAAACAATA  AGCAAGAGAC  AATAATAATG  GCCCTTAATT  ATTAACAAAG  TGCCAGAGTC  2340
TAGGCTAAGC  ACTTTATCTA  TATCTCATTT  CATTCTCACA  ACTTATAAGT  GAATGAGTAA  2400
ACTGAGACTT  AAGGGAACTG  AATCACTTAA  ATGTCACCTG  GCTAACTGAT  GGCAGAGCCA  2460
GAGCTTGAAT  TCATGTTGGT  CTGACATCAA  GGTCTTTGGT  CTTCTCCCTA  CACCAAGTTA  2520
CCTACAAGAA  CAATGACACC  ACACTCTGCC  TGAAGGCTCA  CACCTCATAC  CAGCATACGC  2580
TCACCTTACA  GGGAAATGGG  TTTATCCAGG  ATCATGAGAC  ATTAGGGTAG  ATGAAAGGAG  2640
AGCTTTGCAG  ATAACAAAAT  AGCCTATCCT  TAATAAATCC  TCCACTCTCT  GGAAGGAGAC  2700
TGAGGGGCTT  TGTAAAACAT  TAGTCAGTTG  CTCATTTTTA  TGGGATTGCT  TAGCTGGGCT  2760
GTAAAGATGA  AGGCATCAAA  TAAACTCAAA  GTATTTTTAA  ATTTTTTTGA  TAATAGAAA   2820
ACTTCGCTAA  CCAACTGTTC  TTTCTTGAGT  GTATAGCCCTG  ATCTTGTGGT  AACTTGCTGC  2880
TTCTGCACTT  CATATCCATA  TTTCCTATTG  TTCACTTTAT  TCTGTAGAGC  AGCCTGCCAA  2940
GAATTTTATT  TCTGCTGTTT  TTTTTGCTGC  TAAAGAAAGG  AACTAAGTCA  GGATGTTAAC  3000
AGAAAAGTCC  ACATAACCCT  AGAATTCTTA  GTCAAGGAAT  AATTCAAGTC  AGCCTAGAGA  3060
CCATGTTGAC  TTTCCTCATG  TGTTTCCTTA  TGACTCAGTA  AGTTGGCAAG  GTCCTGACTT  3120
TAGTCTTAAT  AAAACATTGA  ATTGTAGTAA  AGGTTTTTGC  AATAAAAACT  TACTTTGG
```

Seq ID NO: 230 Protein sequence
Protein Accession #:   NP_055213.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MPRQLSAAAA LFASLAVILH DGSQMRAKAP PBTRDYSQPT AAATVQDIKK PVQQPAKQAP   60
HQTLAARFMD GHITFQTAAT VKIPTTTPAT TRNTATTSPI TYTLVTTQAT PNNSHTAPPV  120
TEVTVGPSLA PYSLPPTITP PAHTAGTSSS TVSHTTGNTT QPSNQTTLPA TLSIALEKST  180
TGQKPDQPTH APGTTAAABN TTRTAAPAST VPGPTLAPQP SSVKTGIYQV LNGSRLCIKA  240
EMGIQLIVQD KESVFSPRRY FNIDPNATQA SGNCGTRKSN LLLNFQGGFV NLTFTKDEES  300
YYISEVGAYL TVSDPETVYQ GIKHAVVMFQ TAVGHSFKCV SEQSLQLSAH LQVKTTDVQL  360
QAFDFEDDHP GNVDECSSDY TIVLPVIGAI VVGLCLMGMG VYKIRLRCQS SGYQRI
```

Seq ID NO: 231 DNA sequence
Nucleic Acid Accession #: NM_005409.3
Coding sequence: 94..378

```
1          11         21         31         41         51
|          |          |          |          |          |
TTCCTTTCAT GTTCAGCATT TCTACTCCTT CCAAGAAGAG CAGCAAAGCT GAAGTAGCAG   60
CAACAGCACC AGCAGCAACA GCAAAAAACA AACATGAGTG TGAAGGGCAT GGCTATAGCC  120
TTGGCTGTGA TATTGTGTGC TACAGTTGTT CAAGGCTTCC CCATGTTCAA AAGAGGACGC  180
TGTCTTTGCA TAGGCCCTGG GGTAAAGCA  GTGAAAGTGG CAGATATTGA GAAAGCCTCC  240
ATAATGTACC CAAGTAACAA CTGTGACAAA ATAGAAGTGA TTATTACCCT GAAAGAAAAT  300
AAAGGACAAC GATGCCTAAA TCCCAAATCG AAGCAAGCAA GGCTTATAAT CAAAAAAGTT  360
GAAAGAAAGA ATTTTTAAAA ATATCAAAAC ATATGAAGTC CTGGAAAAGG GCATCTGAAA  420
AACCTAGAAC AAGTTTAACT GTGACTACTG AAATGACAAG AATTCTACAG TAGGAAACTG  480
AGACTTTTCT ATGGTTTTGT GACTTTCAAC TTTTGTACAG TTATGTGAAG GATGAAAGGT  540
GGGTGAAAGG ACCAAAAACA GAAATACAGT CTTCCTGAAT GAATGACAAT CAGAATTCCA  600
CTGCCCAAAG GAGTCCAGCA ATTAAATGGA TTTCTAGGAA AAGCTACCTT AAGAAAGGCT  660
GGTTACCATC GGAGTTTACA AAGTGCTTTC ACGTTCTTAC TTGTTGTATT ATACATTCAT  720
GCATTTCTAG GCTAGAGAAC CTTCTAGATT TGATGCTTAC AACTATTCTG TTGTGACTAT  780
GAGAACATTT CTGTCTCTAG AAGTTATCTG TCTGTATTGA TCTTTATGCT ATATTACTAT  840
CTGTGGTTAC AGTGGAGACA TTGACATTAT TACTGGAGTC AAGCCCTTAT AAGTCAAAAG  900
CATCTATGTG TCGTAAAGCA TTCCTCAAAC ATTTTTTCAT GCAAATACAC ACTTCTTTCC  960
CCAAATATGT TGTAGCACAT CAATATGTAG GGAAACATTC TTATGCATCA TTTGGTTTGT 1020
TTTATAACCA ATTCATTAAA TGTAATTCAT AAAATGTACT ATGAAAAAAA TTATACGCTA 1080
TGGGATACTG GCAACAGTGC ACATATTTCA TAACCAAATT AGCAGCACCG GTCTTAATTT 1140
GATGTTTTTC AACTTTTATT CATTGAGATG TTTTGAAGCA ATTAGGGATT GTGTGTTTAC 1200
TGTACTTTTT GTTTTGATCC GTTTGTATAA ATGATAGCAA TATCTTGGAC ACATTTGAAA 1260
TACAAAATGT TTTTGTCTAC CAAAGAAAAA TGTTGAAAAA TAAGCAAATG TATACCTAGC 1320
AATCACTTTT ACTTTTTGTA ATTCTGTCTC TTAGAAAAAT ACATAATCTA ATCAATTTCT 1380
TTGTTCATGC CTATATACTG TAAAATTTAG GTATACTCAA GACTAGTTTA AAGAATCAAA 1440
GTCATTTTTT TCTCTAATAA ACTACCACAA CCTTTCTTTT TTAAAAAAAA AAA
```

Seq ID NO: 232 Protein sequence
Protein Accession #: NP_005400.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MSVKGMAIAL AVILCATVVQ GFPMFKRGRC LCIGPGVKAV KVADIEKASI MYPSNNCDKI   60
EVIITLKENK GQRCLNPKSK QARLIIKKVE RKNF
```

Seq ID NO: 233 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_000577.1
Coding sequence: 41..520

```
1          11         21         31         41         51
|          |          |          |          |          |
GGCACGAGGG GAAGACCTCC TGTCCTATCA GGCCCTCCCC ATGGCTTTAG AGACGATCTG   60
CCGACCCTCT GGGAGAAAAT CCAGCAAGAT GCAAGCCTTC AGAATCTGGG ATGTTAACCA  120
GAAGACCTTC TATCTGAGGA ACAACCAACT AGTTGCCGGA TACTTGCAAG GACCAAATGT  180
CAATTTAGAA GAAAAGATAG ATGTGGTACC CATTGAGCCT CATGCTCTGT TCTTGGGAAT  240
CCATGGAGGG AAGATGTGCC TGTCCTGTGT CAAGTCTGGT GATGAGACCA GACTCCAGCT  300
GGAGGCAGTT AACATCACTG ACCTGAGCGA GAACAGAAAG CAGGACAAGC GCTTCGCCTT  360
CATCCGCTCA GACAGTGGCC CCACCACCAG TTTTGAGTCT GCCGCCTGCC CCGGTTGGTT  420
CCTCTGCACA GCGATGGAAG CTGACCAGCC CGTCAGCCTC ACCAATATGC CTGACGAAGG  480
CGTCATGGTC ACCAAATTCT ACTTCCAGGA GGACGAGTAG TACTGCCCAG GCCTGCCTGT  540
TCCCATTCTT GCATGGCAAG GACTGCAGGG ACTGCCAGTC CCCCTGCCCC AGGGCTCCCG  600
GCTATGGGGG CACTGAGGAC CAGCCATTGA GGGGTGGACC CTCAGAAGGC GTCACAACAA  660
CCTGGTCACA GGACTCTGCC TCCTCTTCAA CTGACCAGCC TCCATGCTGC CTCCAGAATG  720
GTCTTTCTAA TGTGTGAATC AGAGCACAGC AGCCCTGCA CAAAGCCCTT CCATGTCGCC   780
TCTGCATTCA GGATCAAACC CCGACCACCT GCCCAACCTG CTCTCCTCTT GCCACTGCCT  840
CTTCCTCCCT CATTCCACCT TCCCATGCCC TGGATCCATC AGGCCACTTG ATGACCCCCA  900
ACCAAGTGGC TCCCACACCC TGTTTTACAA AAAAGAAAAG ACCAGTCCAT GAGGGGAGGT  960
TTTAAGGGTT TGTGGAAAAT GAAAATTAGG ATTTCATGAT TTTTTTTTTT CAGTCCCCGT  1020
GAAGGAGAGC CCTTCATTTG GAGATTATGT TCTTTCGGGG AGAGGCTGAG GACTTAAAAT  1080
ATTCCTGCAT TTGTGAAATG ATGGTGAAAG TAAGTGGTAG CTTTTCCCTT CTTTTTCTTC  1140
TTTTTTTGTG ATGTCCCAAC TTGTAAAAAT TAAAAGTTAT GGTACTATGT TAGCCCCATA  1200
ATTTTTTTTT TCCTTTTAAA ACACTTCCAT AATCTGGACT CCTCTGTCCA GGCACTGCTG  1260
CCCAGCCTCC AAGCTCCATC TCCACTCCAG ATTTTTTACA GCTGCCTGCA GTACTTTACC  1320
TCCTATCAGA AGTTTCTCAG CTCCCAAGGC TCTGAGCAAA TGTGGCTCCT GGGGGTTCTT  1380
TCTTCCTCTG CTGAAGGAAT AAAATTGCTCC TTGACATTGT AGAGCTTCTG GCACTTGGAG  1440
ACTTGTATGA AAGATGCTG TGCCTCTGCC TGTCTCCCCC ACCAGGCTGG GAGCTCTGCA  1500
GAGCAGGAAA CATGACTCGT ATATGTCTCA GGTCCCTGCA GGGCCAAGCA CCTAGCCTCG  1560
CTCTTGGCAG GTACTCAGCG AATGAATGCT GTATATGTTG GGTGCAAAGT TCCCTACTTC  1620
CTGTGACTTC AGCTCTGTTT TACAATAAAA TCTTGAAAAT GCCTAAAAAA AAAAAAAAA   1680
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA
```

Seq ID NO: 234 <u>Protein sequence</u>
Protein Accession #: NP_000568.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MALETICRPS GRKSSKMQAF RIWDVNQKTF YLRNNQLVAG YLQGPNVNLE EKIDVVPIEP   60
HALFLGIHGG KMCLSCVKSG DETRLQLEAV NITDLSENRK QDKRFAFIRS DSGPTTSFES  120
AACPGWFLCT AMEADQPVSL TNMPDEGVMV TKFYFQEDE
```

Seq ID NO: 235 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_001840.1
Coding sequence: 149..1567

```
1          11         21         31         41         51
|          |          |          |          |          |
GGGGACTACG GAGAGCTCTG CAGGGAGCCG AGGCCCCCGC CCGGGCCAAG GGAGCTTCTG   60
TCCCGAGGAC CAGGGGATGC GAAGGGATTG CCCCCTGTGG GTCACTTTCT CAGTCATTTT  120
GAGCTCAGCC TAATCAAAGA CTGAGGTTAT GAAGTCGATC CTAGATGGCC TTGCAGATAC  180
CACCTTCCGC ACCATCACCA CTGACCTCCT GTACGTGGGC TCAAATGACA TTCAGTACGA  240
AGACATCAAA GGTGACATGG CATCCAAATT AGGGTACTTC CCACAGAAAT TCCCTTTAAC  300
TTCCTTTAGG GGAAGTCCCT TCCAAGAGAA GATGACTGCG GGAGACAACC CCCAGCTAGT  360
CCCAGCAGAC CAGGTGAACA TTACAGAATT TTACAACAAG TCTCTCTCGT CCTTCAAGGA  420
GAATGAGGAG AACATCCAGT GTGGGGGAGAA CTTCATGGAC ATAGAGTGTT TCATGGTCCT  480
GAACCCCAGC CAGCAGCTGG CCATTGCAGT CCTGTCCCTC ACGCTGGGCA CCTTCACGGT  540
CCTGGAGAAC CTCCTGGTGC TGTGCGTCAT CCTCCACTCC CGCAGCCTCC GCTGCAGGCC  600
TTCCTACCAC TTCATCGGCA GCCTGGCGGT GGCAGACCTG CTGGGGAGTG TCATTTTCGT  660
CTACAGCTTC ATTGACTTCC ACGTGTTCCA CCGCAAAGAT AGCCGCAACG TGTTTCTGTT  720
CAAACTGGGT GGGGTCACGG CCTCCTTCAC TGCCTCCGTG GGCAGCCTGT TCCTCACAGC  780
CATCGACAGG TACATATCCA TTCACAGGCC CCTGGCCTAT AAGAGGATTG TCACCAGGGC  840
CAAGGCCGTG GTGGCGTTTT GCCTGATGTG GACCATAGCC ATTGTGATCG CCGTGCTGCC  900
TCTCCTGGGC TGGAACTGCG AGAAACTGCA ATCTGTTTGC TCAGACATTT TCCCACACAT  960
TGATGAAACC TACCTGATGT TCTGGATCGG GGTCACCAGC GTACTGCTTC TGTTCATCGT  1020
GTATGCGTAC ATGTATATTC TCTGGAAGGC TCACAGCCAC GCCGTCCGCA TGATTCAGCG  1080
TGGCACCCAG AAGAGCATCA TCATCCACAC GTCTGAGGAT GGGAAGGTAC AGGTGACCCG  1140
GCCAGACCAA GCCCGCATGG ACATTAGGTT AGCCAAGACC CTGGTCCTGA TCCTGGTGGT  1200
GTTGATCATC TGCTGGGGCC CTCTGCTTGC AATCATGGTG TATGATGTCT TTGGGAAGAT  1260
GAACAAGCTC ATTAAGACGG TGTTTGCATT CTGCAGTATG CTCTGCCTGC TGAACTCCAC  1320
CGTGAACCCC ATCATCTATG CTCTGAGGAG TAAGGACCTG CGACACGCTT TCCGGAGCAT  1380
GTTTCCCTCT TGTGAAGGCA CTGCGCAGCC TCTGGATAAC AGCATGGGGG ACTCGGACTG  1440
CCTGCACAAA CACGCAAACA ATGCAGCCAG TGTTCACAGG GCCGCAGAAA GCTGCATCAA  1500
GAGCACGGTC AAGATTGCCA AGGTAACCAT GTCTGTGTCC ACAGACACGT CTGCCGAGGC  1560
TCTGTGACGC TGATGCCTCC CTGGCAGCAC AGGAAAGAA TTTTTTTTTT TAAGCTCAAA  1620
ATCTAGAAGA GTCTATTGTC TCCTTGGTTA TATTTTTTTA ACTTTACCAT GCTCAATGAA  1680
AAGGTGATTG CCACATGTCA CTTATTTGCT TAGTTTCCGT TTGGGCTAAT CTTCCGGGGT  1740
TCGTAGGAAA CCTTT
```

Seq ID NO: 236 <u>Protein sequence</u>
Protein Accession #: NP_001831.1

```
      1           11          21          31          41          51
      |           |           |           |           |           |
    MKSILDGLAD  TTFRTITTDL  LYVGSNDIQY  EDIKGDMASK  LGYFPQKFPL  TSFRGSPFQE    60
    KMTAGDNPQL  VPADQVNITE  FYNKSLSSFK  ENEENIQCGE  NPMDIRCFMV  LNPSQQLAIA   120
    VLSLTLGTPT  VLENLLVLCV  ILHSRSLRCR  PSYHFIGSLA  VADLLGSVIF  VYSFIDPHVF   180
    HRKDSRNVPL  FKLGGVTASF  TASVGSLFLT  AIDRYISIHR  PLAYKRIVTR  PKAVVAFCLM   240
    WTIAIVIAVL  PLLGWNCEKL  QSVCSDIFPH  IDETYLMFWI  GVTSVLLLFI  VYAYMYILWK   300
    AHSHAVRMIQ  RGTQKSIIIH  TSEDGKVQVT  RPDQARMDIR  LAKTLVLILV  VLIICWGPLL   360
    AIMVYDVFGK  MNKLIKTVFA  FCSMLCLLNS  TVNPIIYALR  SKDLRHAFRS  MFPSCEGTAQ   420
    PLDNSMGDSD  CLHKHANNAA  SVHRAAESCI  KSTVKIAKVT  MSVSTDTSAE  AL
```

Seq ID NO: 237 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_016083.2
Coding sequence: 64..1482

```
      1           11          21          31          41          51
      |           |           |           |           |           |
    GATTGCCCCC  TGTGGGTCAC  TTTCTCAGTC  ATTTTGAGCT  CAGCCTAATC  AAAGACTGAG    60
    GTTATGAAGT  CGATCCTAGA  TGGCCTTGCA  GATACCACCT  TCCGCACCAT  CACCACTGAC   120
    CTCCTGTACG  TGGGCTCAAA  TGACATTCAG  TACGAAGACA  TCAAAGGTGA  CATGGCATCC   180
    AAAATTAGGGT  ACTTCCCACA  GAAATTCCCT  TTAACTTCCT  TTAGGGGAAG  TCCCTTCCAA   240
    GAGAAGATGA  CTGCGGGAGA  CAACCCCCAG  CTAGTCCCAG  CAGACCAGGT  GAACATTACA   300
    GAATTTTACA  ACAAGTCTCT  CTCGTCCTTC  AAGGAGAATG  AGGAGAACAT  CCAGTGTGGG   360
    GAGAACTTCA  TGGACATAGA  GTGTTTCATG  GTCCTGAACC  CCAGCCAGCA  GCTGGCCATT   420
    GCAGTCCTGT  CCCTCACGCT  GGGCACCTTC  ACGGTCCTGG  AGAACCTCCT  GGTGCTGTGC   480
    GTCATCCTCC  ACTCCCGCAG  CCTCCGCTGC  AGGCCTTCCT  ACCACTTCAT  CGGCAGCCTG   540
    GCGGTGGCAG  ACCTCCTGGG  GAGTGTCATT  TTTGTCTACA  GCTTCATTGA  CTTCCACGTG   600
    TTCCACCGGCA  AAGATAGCCG  CAACGTGTTT  CTGTTCAAAC  TGGGTGGGGT  CACGGCCTCC   660
    CTCACTGCCT  CCGTGGGCAG  CCTGTTCCTC  ACAGCCATCG  ACAGGTACGT  ATCCATTCAC   720
    AGGCCCCTGG  CCTATAAGAG  GATTGTCACC  AGGCCCAAGG  CCGTGGTGGC  GTTTTGCCTG   780
    ATGTGGACCA  TAGCCATTGC  GATCGCCGTG  CTGCCTCTCC  TGGGCTGGAA  CTGCGAGAAA   840
    CTGCAATCTG  TTTGCTCAGA  CATTTTCCCA  CACATTGATG  AAACCTACCT  GATGTTCTGG   900
    ATCGGGGTCA  CCAGCGTACT  GCTTCTGTTC  ATCGTGTATG  CGTACATGTA  TATTCTCTGG   960
    AAGGCTCACA  GCCACGCCGT  CCGCATGATT  CAGCGTGGCA  CCCAGAAGAG  CATCATCATC  1020
    CACACGTCTG  AGGATGGGAA  GGTACAGGTG  ACCCGGCCAG  ACCAAGCCCG  CATGGACATT  1080
    AGGTTAGCCA  AGACCCTGGT  CCTGATCCTG  GTGGTGCTGA  TCATCTGCTG  GGGCCCTTTG  1140
    CTTGCAATCA  TGGTGTATGA  TGTCTTTGGG  AAGATGAACA  AGCTCATTAA  GACGGTGTTT  1200
    GCATTCTGCA  GTATGCTCTG  CCTGCTGAAC  TCCACCGTGA  ACCCCATCAT  CTATGCTCTG  1260
    AGGAGTAAGG  ACCTGCGACA  CGCTTTCCGG  AGCATGTTTC  CCTCTTGTGA  AGGCACTGCG  1320
    CAGCCTCTGG  ATAACAGCAT  GGGGGACTCG  GACTGCCTGC  ACAAACACGC  AAACAATGCA  1380
    GCCAGTGTTC  ACAGGGCCGC  AGAAAGCTGC  ATCAAGAGCA  CGGTCAAGAT  TGCCAAGGTA  1440
    ACCATGTCTG  TGTCCACAGA  CACGTCTGCC  GAGGCCTGCT  GAGCCTGATG  CCTCCCTGGG  1500
    AGCACAGGAA  AAGAATTTTT  TTTTTTAAGC  TCAAAATCTA  GAAGAGTCTA  TTGTCTCCTT  1560
    GGTTATATTT  TTTTAACTTT  ACCATGCTCA  ATGAAAGGT  GATTGTCACC  ATGATCACTT  1620
    ATCAGTTTGC  TAATGTTTCC  ATAGTTTAGG  TACTCAAACT  CCATTCTCCA  GGGGTTTACA  1680
    GTGAAGAAAG  CCTGTTGTTT  AAGTGACTGA  ACGATCCTTC  AAAGTCTCAA  TGAAATAGGA  1740
    GGGAAACCTT  TGGCTACACA  ATTGGAAGTC  TAAGACCCCA  TGGAAAAATG  CCATCAAATG  1800
    AATAATGCCT  TTGTAACCAC  AACTTTCACT  ATAATGTGAA  ATGTAACTGT  CCGTAGTATC  1860
    AGAGATGTCC  ATTTTTACAA  GTTATAGTAC  TAGAGATATT  TTGTAAAATG  TATTATGTCC  1920
    TGTGAGATGT  GTATCAGTGT  TTATGTGCTA  TTAATATTTG  TTTAGTTCAG  CAAAACTGAA  1980
    AGGTAGACTT  TTATGAGAAC  AATGGACAAG  CAGTGGATAC  GTGTCAATGT  GTGCACTTTT  2040
    TTTCTATATT  ATTGCCCATG  ATATAACTTT  AGAAATAAT  CTTAATATTT  CTTCAAATAT  2100
    CTCTATTTAA  TTTTGACACT  GAAATAACCG  TAAAGGTTTA  TTTTTCTGTT  ACCTCAACAA  2160
    GAAGAATTTG  AAGACTTCAA  AATATTGAGC  AGAATTCATT  CATACTTAAA  AATTTATTAG  2220
    CCCTGCATTT  TCATAGGAAG  ACACATTATC  TTCTGGACTA  TAGCTGTTCT  AATGGATTAT  2280
    AATCAGAATG  GAAGAGAGAA  AGCATATTGA  CTTTTTTTGA  GCGGACATCTC  TGACTTTCTT  2340
    TAGTCTTTAG  CTATTACTGG  ATCTCTTAAG  ACAGCATGTG  TTAATCTTAA  TGTATATCGT  2400
    TATCACTGTG  CAGTTGCTGT  TTACTTGAAT  AGTATTGTGT  TCCTATATTC  CAGGTTTAAG  2460
    TAGATTTCAT  GCCTGGGTGG  CCAAACAACA  GTCTTCATTT  TTTTTAATTG  AAAAGAAGTA  2520
    GTGTCTGGAT  CAGTAAAATT  ATACTGTGTG  TGAGTGTGAA  TATAAATGTG  TGTATGTGTG  2580
    TTTCTGTCCG  TAACTGTTAC  AGTAATGTCA  TAAAGTGAGA  AAACTGTGAC  CAAGTATAAA  2640
    CTTTTACCAC  TTGCTGCACT  CTTGCACATG  GATTCAGTTT  CTAAAATTGA  GTTCTTCCTG  2700
    TAATCTTGTT  GATAAAAATA  CTGACTCCAA  CCATTCAAAA  ATTTCACCCC  ATCCCTCCTT  2760
    AAGAGATTGG  ATCAAGTATT  ACTAAATTGA  CCTTTAGGTA  TTACACAAGA  CCAGTGCTTA  2820
    GCAAAAAATA  ATGACAGGCA  TCCAAGGAAG  GGATGTATTT  GTAGTGTTAT  TGCCAGGAAA  2880
    GGAGAGTACT  TTGGTTTCTG  AGCACCGAAT  ATTGAGCAAT  ATGTCAGTCA  CTAAAAGGAA  2940
    GACAGTTCTA  CAGAAAAACA  AATGGTAACA  TTTTTCAATA  GCGTGTGTGA  ATAGTATGCA  3000
    CTATATACAT  CACGTTAAAG  TAGGACTATC  ACACCCAGCC  CATGTGGCTA  AAAAAGCTGA  3060
    ATCAGACAGT  GGATGAGACA  CACAACGGCA  GTGAAGAACC  GATACACTTG  GCATTGACGT  3120
    CTAGCTATGC  TGTATCTGTG  CTTTGCCCAC  ATGCCCTTGG  TGACAGCTGA  GCACCCAGCT  3180
    CTGTCTTGGT  AGGTTTGGGC  TAAGGAACAA  ATCTCTCCTT  TGCTCGGTGGT  TAGCAAGATA  3240
    CACTCAAGCA  TGAAGATAAA  CACAGCTGCT  TTCTTCTTAC  ACCCCGGTCT  CATGCTCCTT  3300
    AATGGCGCCA  TGGGTGCTTG  TTGGGCCTTT  TTCCAGTAAG  GAATGATATT  GCTGAAGAAT  3360
    CTACTTAACC  CTGACAAATT  TTAATTATAA  TCTCTTCTTA  TACAGATAAA  ACATGACTCC  3420
    TACAAGGCCC  CAAGGTTTAC  ATAGTCTGAA  GTGAAGTACA  GAGCTGGCAT  CTATCTGGTG  3480
    ATTTCTAGCT  CTCGAGATAC  CCAAGCAGCC  TGATGGGGCA  GTTCCCCTTC  TTACGGTTCA  3540
    CGCTCTAAGG  CAGGATGTGG  CTTATGAGAT  ACTTTGCATT  GTCTGTCTGC  ACACCTTGAA  3600
    TCTGCCTGCT  GGCTCCCTTA  CTTTACCTCT  CTGTCATGTG  CAGATGAAGG  CTCAGGGTGC  3660
    TAGAGGATTA  GTAAGATCTC  TTTCTAAAGA  CAGGAGAGAT  TATTTACAAG  AAGAACTCAC  3720
    CAGGGTTTAG  TTTGCATTTA  AGAATTGCCA  GTCTTTGTC  CTGCTCGGTGGT  TTGAACATTA  3780
    ATCCACATGT  TTCGAGCTC  ACCAGGCAGT  ACCAATGCTC  TTTTCACAGC  TATGAAGAGC  3840
    TAGAGAAATT  CTTGTTATGG  TAGAAAAATT  TCACGGTTCA  TTTTTGAAAC  TGCATTTGTG  3900
    CGTATGCAGT  GTAGATTTTA  TAGTGTGTTG  TGCTTTCAAG  ATCTAAATCA  TATATAATAA  3960
```

```
            ATTAAGGGAC AATGGGGCTG ACAGCACTAA ACTTGGTGCT TATTGATATT CTAAGAAATA  4020
            TCTGTGAAAT ATCATCACGT ATGTTATACA ACCTTCATTT AAAAAGGTTT AAAACTAGTT  4080
            AGATTCACTT TGACACTTTT CATATCATTT CTTAACCCAA GTGACGAAAA CATTGTCCCC  4140
            AATGAATATA CTCATTAGAA TTACCATTTG TTAATATCAC TCATTAATTA ACCCCATAAT  4200
 5          TAGATCCATT AAATTTAAATG ATTTAAATTT AAGTAAGTTT TATAAGGTCT GACATCAGAG  4260
            GTATCTTACT TTCCTCTGAG GATGATGTAC TTGCCCTGAC CATGCATTT ACCATCACAC  4320
            ATGTTCAGAA AGGGCCAAAT TCCCAACCTG CTCATTTTTT TTTTTATCAG AGTCATGATG  4380
            AATCAGTCCT AGAATGTTTC ATTTGCACAA GTAGGGCTGC CTCCAAGAGG AACCTCTGAT  4440
            TTATTTTGTA TGAAATATAT GTGAAAGGAT ATGAATCTGA GAGATGCTGT AGACATCTGT  4500
10          CCTACACTTG AGATGATTTC CAAGCCTCTC TGGCACTTTG AGTTAAGTCT ATCTGGTATT  4560
            AAATGCCAAG GACCTTTTGC TGCCTAAATC CACTCTGCAG GAAATAGGCC CAACCACCAG  4620
            ATGAGAATTA GGCCCTGGAT GAGTAGCCGCT ATAGTTACTG TCCTGTTGAT TAATTTCTGC  4680
            CATTTCATGT CCATAAAAGA GACCACCCAT ATCATGCACA CAATTAGATT TCTCACACTC  4740
15          TAACTGTATA TTTGTATGAT ATTTTAAAAT CTCCTAAATG CTGGGCAATG GCTATTAACA  4800
            ATTAATTGTC TTGCACTGGC CTTCTGATGA AATGTTAACA ATGCCTATTG TAATATAGAA  4860
            AAAAACATTC TATCTACTGA TTTGGGCTGA ATGTATGTAA ATAGGTTTCT AAAAAGTCAG  4920
            ATGTTGAGC AGTGGCCTAC AAATCAGTAA TTTTCGGGTG GGAGAGTTTC TTTACATTGC  4980
            CGTGGCATCT TAAAAGCTAT CTTCATGTAA ATTGACTGTA CTAGGCCTAC TGGGGATCGA  5040
20          AGTTCCCAAG AAAGGAAACC TTTTCTTGTA TCTGGATTCA AATTTATTTC CAATGTTTCA  5100
            AGCGGGAAAC ATGACTCTTT ATTGTCTGTA AATCTAACAT TATTACTTTT CCTCTTAGAA  5160
            GAATATTGTA TTGTTAGATG TTTGTTGAGC TGGTAACATC GTTGCAACCA CTGCAATATC  5220
            TTCGTTAGTA ATCTGTATAA TACTTTGTAT ACAAGTACTG GTAAGATTGT TATTAAATGT  5280
            AGCTTCAGTC ATTAAATTAC TATAGCAAAG TAGTACTTCT TCTGTAATAT TTACAATGTA  5340
25          TTAAGCCCAC AGTATATTTT ATTTCAATGT AATTAAACTG TTAACTTATT CAAAGAGAAA  5400
            ACATCTCATC ATGTCTATTG TCCAAAGTTA CCTGGAATCA AATAAAAATT CTAGATTACC  5460
            ATGAAGAACA TA
```

Seq ID NO: 238 Protein sequence
Protein Accession #: NP_057167.1

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          MKSILDGLAD TTFRTITTDL LYVGSNDIQY EDIKGDMASK LGYFPQKFPL TSFRGSPFQE   60
          KMTAGDNPQL VPADQVNITE FYNKSLSSFK ENEENIQCGE NFMDIECFMV LNPSQQLAIA  120
          VLSLTLGTFT VLENLLVLCV ILHSRSLRCR PSYHFIGSLA VADLLGSVIF VYSFIDFHVF  180
          HRKDSRNVFL FKLGGVTASL TASVGSLFLT AIDRYVSIHR PLAYKRIVTR PKAVVAFCLM  240
          WTIAIAIAVL PLLGWNCEKL QSVCSDIFPH IDETYLMFWI GVTSVLLLFI VYAYMYILWK  300
          AHSHAVRMIQ RGTQKSIIIH TSEDGKVQVT RPDQARMDIR LAKTLVLILV VLIICWGPLL  360
          AIMVYDVFGK MNKLIKTVPA FCSMLCLLNS TVNPIIYALR SKDLRHAFRS MFPSCEGTAQ  420
          PLDNSMGDSD CLHKHANNAA SVHRAAESCI KSTVKIAKVT MSVSTDTSAE AL
```

Seq ID NO: 239 DNA sequence
Nucleic Acid Accession #: NM_033181.1
Coding sequence: 17..1252

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          ATGAAGTCGA TCCTAGATGG CCTTGCAGAT ACCACCTTCC GCACCATCAC CACTGACCTC   60
          CTGTACGTGG GCTCAAATGA CATTCAGTAC GAAGACATCA AAGGAGAATG AGGAGAACAT  120
          CCAGTGTGGG GAGAACTTCA TGGACATAGA GTGTTTCATG GTCCTGAACC CCAGCCAGCA  180
          GCTGGCCATT GCAGTCCTGT CCCTCACGCT GGGCACCTTC GCGGTCCTGG AGAACCTCCT  240
          GGTGCTGTGC GTCATCCTCC ACTCCCGCAG CCTCCGCTGC AGGCCTTCCT ACCACTTCAT  300
          CGGCAGCCTG GCGGTGGCAG ACCTCCTGGG GAGTGTCATT TTTGTCTACA GCTTCATTGA  360
          CTTCCACGTG TTCCACCGCA AAGATAGCCG CAACGTGTTT CTGTTCAAAC TGGGTGGGGT  420
          CACGGCCTCC TTCACTGCCT CCGTGGGCAG CCTGTTCCTC ACAGCCATCG ACAGGTACAT  480
          ATCCATTCAC AGGCCCCTGG CCTATAAGAG GATTGTCACC AGGCCCAAGG CCGTGGTGGCA  540
          GTTTTGCCTG ATGTGGACCA TAGCCATTGT GATCGCCGTG CTGCCTCTCC TGGGCTGGAA  600
          CTGCGAGAAA CTGCAATCTG TTTGCTCAGA CATTTTCCCA CACATTGATG AAACCTACCT  660
          GATGTTCTGG ATCGGGGTCA CCAGCGTACT GCTTCTGTTC ATCGTGTATG CGTACATGTA  720
          TATTCTCTGG AAGGCTCACA GCCACGCCGT CCGCATGATT CAGCGTGGCA CCCAGAAGAG  780
          CATCATCATC CACACGTCTG AGGATGGGAA GGTACAGGTG ACCCGGCCAG ACCAAGCCCG  840
          CATGGACATT AGGTTAGCCA AGACCCTGGT CCTGATCCTG GTGGTGTTGA TCATCTGCTG  900
          GGGCCCTCTG CTTGCAATCA TGGTGTATGA TGTCTTCATA GGGAAGATGA ACAAGCTCAT  960
          GACGGTGTTT GCATTCTGCA GTATGCTCTG CCTGCTGAAC TCCACCGTGA ACCCCATCAT  1020
          CTATGCTCTG AGGAGTAAGG ACCTGCGACA CGCTTTCCGG AGCATGTTTC CCTCTTGTGA  1080
          AGGCACTGCG CAGCCTCTGG ATAACAGCAT GGGGGACTCG GACTGCCTGC ACAAACACGC  1140
          AAACAATGCA GCCAGTGTTC ACAGGGCCGC AGAAAGCTGC ATCAAGAGCA CGGTCAAGAT  1200
          TGCCAAGGTA ACCATGTCTG TGTCCACAGA CACGTCTGCC GAGGCTCTGT GA
```

Seq ID NO: 240 Protein sequence
Protein Accession #: NP_149421.1

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          MALQIPPSAP SPLTSCIWAQ MTFSTKTSKE NEENIQCGEN FMDIECFMVL NPSQQLAIAV   60
          LSLTLGTFIV LENLLVLCVI LHSRSLRCRP SYHFIGSLAV ADLLGSVIFV YSFIDFHVFH  120
          RKDSRNVFLF KLGGVTASFT ASVGSLFLTA IDRYISIHRP LAYKRIVTRP KAVVAFCLMW  180
          TIAIVIAVLP LLGWNCEKLQ SVCSDIFPEI DETYLMFWIG VTSVLLLFIV YAYMYILWKA  240
          HSHAVRMIQR GTQKSIIIHT SEDGKVQVTR PDQARMDIRL AKTLVLILVV LIICWGPLLA  300
          IMVYDVFGRM NKLIKTVPAF CSMLCLLNST VNPIIYALRS KDLRHAFRSM FPSCEGTAQP  360
          LDNSMGDSDC LHKHANNAAS VHRAAESCIK STVKIAKVTM SVSTDTSAEA L
```

Seq ID NO: 241 DNA sequence
Nucleic Acid Accession #: NM_003596.1

Coding sequence: 82..1194

```
      1         11        21        31        41        51
      |         |         |         |         |         |
GTAGACTGTC CATGGCCTGA ACATTTTCCG AAAATCATTT TGAGCAAAAT ATCTGTTTAA   60
TAACAAGATA ACCACATCAA GATGGTTGGA AAGCTGAAGC AGAACTTACT ATTGGCATGT  120
CTGGTGATTA GTTCTGTGAC TGTGTTTTAC CTGGGCCAGC ATGCCATGGA ATGCCATCAC  180
CGGATAGAGG AACGTAGCCA GCCAGTCAAA TTGGAGAGCA CAAGGACCAC TGTGAGAACT  240
GGGCTGGACC TCAAAGCCAA CAAAACCTTT GCCTATCACA AAGATATGCC TTTAATATTT  300
ATTGGRGGTG TGCCTCGGAG TGGAACCACA CTCATGAGGG CCATGCTGGA CGCACATCCT  360
GACATTCGCT GTGGAGAGGA AACCAGGGTC ATTCCCCGAA TCCTGGCCCT GAAGCAGATG  420
TGGTCACGGT CAAGTAAAGA GAAGATCCGC CTGGATGAGG CTGGTGTTAC TGATGAAGTG  480
CTGGATTCTG CCATGCAAGC CTTCTTACTA GAAATTATCG TTAAGCATGG GGAGCCAGCC  540
CCTTATTTAT GTAATAAAGA TCCTTTTGCC CTGAAATCTT TAACTTACCT TTCTAGGTTA  600
TTCCCCAATG CCAAATTTCT CCTGATGGTC CGAGATGGCC GGGCATCAGT ACATTCAATG  660
ATTTCTCGAA AAGTTACTAT AGCTGGGATT GATCTGAACA GCTATAGGGA CTGTTTGACA  720
AAGTGGAATC GTGCTATAGA GACCATGTAT AACCAGTGTA TGGAGGTTGG TTATAAAAAG  780
TGCATGTTGG TTCACTATGA ACAACTTGTC TTACATCCTG AACGGTGGAT GAGAACACTC  840
TTAAAGTTCC TCCAGATTCC ATGGAACCAC TCAGTATTGC ACCATGAAGA GATGGATTGGG  900
AAAGCTGGGG GAGTGTCTCT GTCAAAAGTG GAGAGATCTA CAGACCAAGT AATCAAGCCA  960
GTCAATGTAG GAGCTCTATC AAAATGGGGT GGGAAGATAC CGGCAGATGT TTTACAAGAC 1020
ATGGCAGTGA TTGCTCCTAT GCTTGCCAAG CTTGGATATG ACCCATATGC CAACCCACCT 1080
AACTACGGAA AACCTGATCC CAAAATTATT GAAAACACTC GAAGGGTCTA TAAGGGAGAA 1140
TTCCAACTAC CTGACTTTCT TAAAGAAAAA CCACAGACTG AGCAAGTGGA GTAGCAGAAC 1200
CAGGAGCCTC TTCCATACAT GAGGAAAGAT TGCTGCCTTT TCAGCAGAAG GGAAATTCCT 1260
AGGATTGGCT GTCCCCTGCC AAGCTTGGTG GAGCGTCTGC ACCTTGGCTG CGCCGCCTGT 1320
GCATTTGCCA GTTTCCTCCC ACTGAGAGGA TGGAGGTGGA CGCACAGCTT TGGGCCTCGT 1380
GAGGGATCTG CCTCCTGAGC AAAGAGCTCT TGATCCCGAT TTCATGCACA GCCCTGCAGT 1440
AAGGAGCCCA GAAGGAACAT GTGTTTCCTG TTAAAACTCC TCTTGTTCTC TTTTCTTACA 1500
TTATGACGTT TGTTTTCAAG GAGAGGGTTT AAAAATGGGA TCCTGTAAGC AGACTTGGGC 1560
AGTCTCCTTT TGAAATAGGT TGTCTGTACA TGTTCTAATG TTTTGTAGAA CACGTGTGCC 1620
TGTTTAAGTG TATTGATGTG AATAATATTA AATATCCTAA TTATTTAATT CATTGTATTG 1680
TTTCTGAGAA GTTGGGAAAT TACCATTATA CATTTACAAC CTAATGACTT TTGTATTTTA 1740
TTTTTCAAAA TAAAAGCTTT CAATGTGA
```

Seq ID NO: 242 Protein sequence
Protein Accession #: NP_003587.1

```
      1         11        21        31        41        51
      |         |         |         |         |         |
MVGKLKQNLL LACLVISSVT VPYLGQHAME CHHRIEERSQ PVKLESTRTT VRTGLDLKAN   60
KTFAYHKDMP LIFIGGVPRS GTTLMRAMLD AHPDIRCGEE TRVIPRILAL KQMWSRSSKE  120
KIRLDEAGVT DEVLDSAMQA FLLEIIVKHG EPAPYLCNKD PFALKSLTYL SRLFPNAKFL  180
LMVRDGRASV HSMISRKVTI AGFDLNSYRD CLTKWNRAIE TMYNQCMEVG YKKCMLVHYE  240
QLVLHPERWM RTLLKFLQIP WNHSVLHHEE MIGKAGGVSL SKVERSTDQV IKPVNVGALS  300
KWVGKIPPDV LQDMAVIAPM LAKLGYDPYA NPPNYGKPDP KIIENTRRVY KGEFQLPDFL  360
KEKPQTEQVE
```

Seq ID NO: 243 DNA sequence
Nucleic Acid Accession #: NM_001492.3
Coding sequence: 1395..2513

```
      1         11        21        31        41        51
      |         |         |         |         |         |
ACGCGGGGCG CGCGGCTCCG TCGGCTACCG CGGGCGGGCG CAGGCGACGG GCACGGCGGG   60
CGAGCGGGCG GTATGGCGGC GGCGGGGCCC GCGGCGGGGC CGACGGGGCC CGAGCCCATG  120
CCGAGCTACG CGCAGCTAGT GCAGCGCGGC TGGGGCAGCG CGCTGGCGGC GGCGCGGGGC  180
TGCACGGACT GCGGCTGGGG GCTGGCGCGT CGCGGCCTGG CTGAGCACGC GCACCTGGCG  240
CCGCCCGAGC TGCTGCTGCT GGCGCTCGGC GCGCTGGGCT GGACGCGCCT GCGCTCCGCG  300
GCCACTGCGC GCCTCTTTCG GCCCCTGGCG AAGCGGTGCT GCCTCCAGCC CAGAGATGCC  360
GCCAAGATGC CCGAGAGCGC TTGGAAGTTT CTCTTCTACC TGGGCAGCTG GAGCTACAGT  420
GCCTACCTGC TGTTTGGCAC CGACTACCCC TTCTTCCATG ACCCACCATC TGTCTTCTAC  480
GACTGGACGC CGGGCATGGC AGTGCCACGG GACATTGCAG CCGCCTACCT GCTCCAGGGA  540
AGCTTCTATG GCCACTCCAT CTACGCTACG CTATACATGG ACACCTGGCG CAAGGACTCG  600
GTGGTCATGC TGCTCCACCA CGTGGTCACT CTCATCCTCA TCGTCTCCTC CTACGCCTTC  660
CGGTACCACA ATGTGGGCAT CCTTGTGCTC TTCCTGCACG ATATCAGTGA CGTGCAGCTT  720
GAGTTCACCA AGCTCAACAT TTACTTCAAG TCCCGCGGCG GCTCCTACCA TCGGCTGCAT  780
GCCTTGGCAG CAGACTTGGG CTGCCTCAGC TTCGGCTTCA GCTGGTTCTG GTTCCGGCTC  840
TACTGGTTCC CGCTCAAGGT CCTGTATGCC ACCAGTCACT GCAGTCTGCG CACGGTGCCT  900
GACATCCCCT TCTACTTCTT CTTCAATGCG CTCCTGCTGC TGCTCACCCT TATGAACCTC  960
TACTGGTTCC TGTACATCGT GGCGTTTGCA GCCAAGGTGT TGACAGGCCA GGTGCACGAG 1020
CTGAAGGACC TGCGGGAGTA TGACACAGCC GAGGCCCAGA GCCTGAAGCC CAGCAAAGCC 1080
GAGAAGCCAC TGAGGAACGG CCTGGTGAAG GACAAGCGCT TCTGAACCCC TCGGCCCCGC 1140
CCCCGTGGAC CCGGCCCCAC CCCGAATACC CCGGCCACGC TCCCCGTCCT TGGCCGCCCC 1200
TCCACCCCCT CCAACTCTGC TCCTCTAGGG CCGCCGCCAC CTCCCCTGGG ACCCCGCCCC 1260
CTCATCCTGC CTCCATTCCC CGGCCACGCC CCCCAGGACC CCTGCCCCTC CGGGGACACC 1320
GGCCCCGCCC TCAGCCCACT GGTCCCGGGC CGCCGCGGAC CCTGCGCACT CTCTGGTCAT 1380
CGCCTGGGAG GAAGATGCCA CCGCCGCAGC AAGGTCCCTG CGGCCACCAC CTCCTCCTCC 1440
TCCTGGCCCT GCTGCTGCCC TCGCTGCCCC TGACCGCCGC CCCCGTGCCC CCAGGCCCAG 1500
CCGCCGCCCT GCTCCAGGCT CTAGGACTGC GCGATGAGCC CCAGGGTGCC CCCAGGCTCC 1560
GGCCGGTTCC CCCGGTCATG TGGCGCCTGT TTCGACGCGG GGACCCCCAG GAGACCAGGT 1620
CTGGCTCGCG GCGGACGTCC CCAGGGGTCA CCCTGCAACC GTGCCACGTG GAGGAGCTGG 1680
GGGTCGCCGG AAACATCGTG CGCCACATCC CGGACCGCGG TGCGCCCACC CGGGCCTCGG 1740
AGCCTGTCTC GGCCGCGGGG CATTGCCCTG AGTGGACAGT CGTCTTCGAC CTGTCGGCTG 1800
```

```
TGGAACCCGC TGAGCGCCCG AGCCGGGCCC GCCTGGAGCT GCGTTTCGCG GCGGCGGCGG   1860
CGGCAGCCCC GGAGGGCGGC TGGGAGCTGA GCGTGGCGCA AGCGGGCCAG GGCGGCGGGCG   1920
CGGACCCCGG GCCGGTGCTG CTCCGCCAGT TGGTGCCCGC CCTGGGGCCG CCAGTGCGCG   1980
CGGAGCTGCT GGGCGCCGCT TGGGCTCGCA ACGCCTCATG GCCGCGCAGC CTCCGCCTGG   2040
CGCTGGCGCT ACGCCCCCGG GCCCCTGCCG CCTGCGCGCG CCTGGCCGAG GCCTCGCTGC   2100
TGCTGGTGAC CCTCGACCCG CGCCTGTGCC ACCCCCTGGC CCGGCCGCGG CGCGACGCCG   2160
AACCCGTGTT GGGCGGCGGC CCCGGGGGCG CTTGTCGCGC GCGGCGGCTG TACGTGAGCT   2220
TCCGCGAGGT GGGCTGGCAC CGCTGGGTCA TCGCGCCGCG CGGCTTCCTG GCCAACTACT   2280
GCCAGGGTCA GTGCGCGCTG CCCGTCGCGC TGTCGGGGTC CGGGGGCCG CCGGCGCTCA   2340
ACCACGCTGT GCTGCGCGCG CTCATGCACG CGGCCGCCCC GGGAGCGGCC GACCTGCCCT   2400
GCTGCGTGCC CGCGCGCCTG TCGCCCATCT CCGTGCTCTT CTTTGACAAC AGCGACAACG   2460
TGGTGCTGCG GCAGTATGAG GACATGGTGG TGGACGAGTG CGGCTGCCGC TAACCCGGGG   2520
CGGGCAGGGA CGCGGGCCCA ACAATAAATG CCGCGTGGTC TGCTC
```

Seq ID NO: 244 Protein sequence
Protein Accession #: NP_001483.2

```
1          11         21         31         41         51
|          |          |          |          |          |
MPPPQQGPCG HHLLLLLALL LPSLPLTRAP VPPGPAAALL QALGLRDEPQ GAPRLRFVPP   60
VMWRLFRRRD PQETRSGSRR TSPGVTLQPC HVEELGVAGN IVRHIPDRGA PTRASEPVSA   120
AGHCPEWTVV FDLSAVEPAE RPSRARLELR PAAAAAAPE GGWELSVAQA GQGAGADPGP   180
VLLRQLVPAL GPPVRAELLG AAWARNASWP RSLRLALALR PRAPAACARL AEASLLLVTL   240
DPRLCHPLAR PRRDAEPVLG GGPGGACRAR RLYVSFREVG WHRWVIAPRG FLANYCQGQC   300
ALPVALSGSG GPPALNHAVL RALMHAAAPG AADLPCCVPA RLSPISVLFF DNSDNVVLRQ   360
YEDMVVDECG CR
```

Seq ID NO: 245 DNA sequence
Nucleic Acid Accession #: NM_021267.1
Coding sequence: 17..1125

```
1          11         21         31         41         51
|          |          |          |          |          |
ACGCGGGCGG CGCGGCTCCG TCGGCTACCG CGGGCGGGCG CAGGCGACGG GCACGGCGGG   60
CGAGCGGGCG GTATGGCGGC GGCGGGGCCC GCGGCGGGGC CGACGGGGCC CGAGCCCATG   120
CCGAGCTACG CGCAGCTAGT GCAGCGCGGC TGGGGCAGCG CGCTGGCGGC GGCGCGGGGC   180
TGCACGGACT GCGGCTGCGGG GCTGGCGCGT CGCGGCCTGG CTGAGCACGC GCACCTGGCG   240
CCGCCCGAGC TGCTGCTGCT GGCGCTGGCG GCGCTGGGCT GGACCGCGCT GCGCTCCGCG   300
GCCACTGCGC GCCTCTTTCG GCCCCTGGCG AAGCGGTGCT GCTCCAGCC CAGAGGATGCC   360
GCCAAGATGC CCGAGAGCGC TTGGAAGTTT CTCTTCCTACC TGGGCAGCTG GAGCTACAGT   420
GCCTACCTGC TGTTTGGCAC CGACTACCCG TTCTTCCATG ACCCACCATC TGTCTTCTAC   480
GACTGGACGC CGGGCATGGC AGTGCCACGG GACATTGCAG CCGCCTACCT GCTCCAGGGA   540
AGCTTCTATG GCCACTCCAT CTACGCTACG CTATACATGG ACACCTGGCG CAAGGACTCG   600
GTGGTCATGC TGCTCCACCA CGTGGTCACT CTCATCCTCA TCGTCTCCTC CTACGCCCTC   660
CGGTACCACA ATGTGGGCAT CCTTGTGCTC TTCCTGCACG ATATCAGTGA CGTGCAGCTT   720
GAGTTCACCA AGCTCAACAT TTACTTCAAG TCCCGCGGCG GCTCCTACCA TCGGCTGCAT   780
GCCTTGGCAG CAGACTTGGG CTGCCTCAGC TTCGGCCTCA GTTCCGCCTC GTTCCGGCCTC   840
TACTGGTTCC CGCTCAAGGT CCTGTATGCC ACCAGTCACT GCAGTCTGCG CACGGTGCCT   900
GACATCCCCT TCTACTTCTT CTTCAATGCG CTCCTGCTGC TGCTCACCCT TATGAACCTC   960
TACTGGTTCC TGTACATCGT GGCGTTTGCA GCCAAGGTGT TGACAGGCCA GGTGCACGAG   1020
CTGAAGGACC TGCGGGAGTA TGACACAGCC GAGGCCCAGA GCCTGAAGCC CAGCAAAGCC   1080
GAGAAGCCAC TGAGGAACGG CCTGGTGAAG GACAAGCGCT TCTGAACCCC TCGGCCCCGC   1140
CCCCGTGGAC CCGGCCCCAC CCCGAATACC CCGGCCACGC TCCCCGTCCT TGGCCGCCCC   1200
TCCACCCCCT CCAACTCTGC TCCTCTAGGG CCGCCGCCAC CTCCCTGGG ACCCCGCCCC   1260
CTCATCCTGC CTCCATTTCC CGGCCACGCC CCCCAGGACC CCTGCCCCTC CGGGGACACC   1320
GGCCCCGCCC TCAGCCCACT GGTCCCCGGC CGCCGCGGAC CCTGCGCACT CTCTGGTCAT   1380
CGCCTGGGAG GAAGATGCCA CCGCCGCAGC AAGGTCCCTG CGGCCACCAC CTCCTCCTCC   1440
TCCTGCCCCT GCTGCTGCCC TCGCTGCCCC TGACCCGCGC CCCCGTGCCC CCAGGCCCAG   1500
CCGCCGCCCT GCTCCAGGCT CTAGGACTGC GCGATGAGCC CCAGGGTGCC CCCAGGCTCC   1560
GGCCGGTTCC CCCGGTCATG TGGCGCCTGT TTCGACGCGG GGACCCCCAG GAGACCAGGT   1620
CTGGCTCGCG GCGGACGTCC CCAGGGGTCA CCCTGCAACC GTGCCACGTG GAGGAGCTGG   1680
GGGTCGCCGG AAACATCGTG CGCCACATCC CGGACCGCGG TGCGCCCACC CGGGCCTCGG   1740
AGCCTGTCTC GGCCGCGGGG CATTGCCCTG AGTGGACAGT CGTCTTCGAC CTGTCGGCTG   1800
TGGAACCCGC TGAGCGCCCG AGCCGGGCCC GCCTGGAGCT GCGTTTCGCG GCGGCGGCGG   1860
CGGCAGCCCC GGAGGGCGGC TGGGAGCTGA GCGTGGCGCA AGCGGGCCAG GGCGGCGGGCG   1920
CGGACCCCGG GCCGGTGCTG CTCCGCCAGT TGGTGCCCGC CCTGGGGCCG CCAGTGCGCG   1980
CGGAGCTGCT GGGCGCCGCT TGGGCTCGCA ACGCCTCATG GCCGCGCAGC CTCCGCCTGG   2040
CGCTGGCGCT ACGCCCCCGG GCCCCTGCCG CCTGCGCGCG CCTGGCCGAG GCCTCGCTGC   2100
TGCTGGTGAC CCTCGACCCG CGCCTGTGCC ACCCCCTGGC CCGGCCGCGG CGCGACGCCG   2160
AACCCGTGTT GGGCGGCGGC CCCGGGGGCG CTTGTCGCGC GCGGCGGCTG TACGTGAGCT   2220
TCCGCGAGGT GGGCTGGCAC CGCTGGGTCA TCGCGCCGCG CGGCTTCCTG GCCAACTACT   2280
GCCAGGGTCA GTGCGCGCTG CCCGTCGCGC TGTCGGGGTC CGGGGGCCG CCGGCGCTCA   2340
ACCACGCTGT GCTGCGCGCG CTCATGCACG CGGCCGCCCC GGGAGCGGCC GACCTGCCCT   2400
GCTGCGTGCC CGCGCGCCTG TCGCCCATCT CCGTGCTCTT CTTTGACAAC AGCGACAACG   2460
TGGTGCTGCG GCAGTATGAG GACATGGTGG TGGACGAGTG CGGCTGCCGC TAACCCGGGG
```

Seq ID NO: 246 Protein sequence
Protein Accession #: NP_067090.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MAAAGPAAGP TGPEPMPSYA QLVQRGWGSA LAAARGCTDC GWGLARRGLA EHAHLAPPEL   60
LLLALGALGW TALRSAATAR LFRPLAKRCC LQPRDAAKMP ESAWKFLFYL GSWSYSAYLL   120
FGTDYPFFHD PPSVFYDWTP GMAVPRDIAA AYLLQGSFYG HSIYATLYMD TWRKDSVVML   180
```

1046

LHHVVTLILI VSSYAFRYHN VGILVLFLHD ISDVQLEFTK LNIYFKSRGG SYHRLHALAA 240
DLGCLSFGFS WFWFRLYWFP LKVLYATSHC SLRTVPDIPF YFFFNALLLL LTLMNLYWFL 300
YIVAPAAKVL TGQVHELKDL REYDTAEAQS LKPSKAEKPL RNGLVKDKRF

Seq ID NO: 247 DNA sequence
Nucleic Acid Accession #: NM_002081.1
Coding sequence: 222..1898

```
        1        11        21        31        41        51
        |         |         |         |         |         |
        |         |         |         |         |         |
GGCTGCCCGA GCGAGCGTTC GGACCTCGCA CCCCGCGCGC CCCGCGCCGC CGCCGCCGCC   60
GGCTTTTGTT GTCTCCGCCT CCTCGGCCGC CGCCGCGCGG GGACCGCGAG CCGCGCGCGC  120
CGGGACCTTG GCTCTGCCCT TCGCGGGCGG GAACTGCGCA GGACCCGGCC AGGATCCGAG  180
AGAGGCGCGG GCGGGTGGCC GGGGGCGCCG CCGGCCCCGC CATGGAGCTC CGGGCCCGAG  240
GCTGGTGGCT GCTATGTGCG GCCGCAGCGG TGGTCGCCGG CGCCCGCGGG GACCCGGCCA  300
GCAAGAGCCG GAGCTGCGGC GAGGTCCGCC AGATCTACGG AGCCAAGGGC TTCAGCCTGA  360
GCGACGTGCC CCAGGCGGAG ATCTCGGGTG AGCACCTGCG GATCTGTCCC CAGGGCTACA  420
CCTGCTGCAC CAGCGAGATG GAGGAGAACC TGGCCAACCG CAGCCATGCC GAGCTGGAGA  480
CGGCGCTCCG GGACAGCAGC CGCGTCCTGC AGGCCATGCT TGCCACCCAG CTGCGCAGCT  540
TCGATGACCA CTTCCAGCAC CTGCTGAACG ACTCGGAGCG GACGCTGCAG GCCACCTTCC  600
CCGGCGCCTT CGGAGAGCTG TACACGCAGA ACGCGAGGGC CTTCCGGGAC CTGTACTCAG  660
AGCTGCGCCT GTATACCGC GGTGCCAACC TGCACCTGGA GGAGACGCTG GCCGAGTTCT  720
GGGCCCGCCT GCTCGAGCGC CTCTTCAAGC AGCTGCACCC CCAGCTGCTG CTGCCTGATG  780
ACTACCTGGA CTGCCTGGGC AAGCAGGCCG AGGCGCTGCG GCCCTTCGGG GAGGCCCCGA  840
GAGAGCTGCG CCTGCGGGCC ACCCGTGCCT TCGTGGGGCA CTCGCTCCTT GTGCGAGGCG  900
TGGGCGTGGC CAGCGGACGTG GTCCGGAAAG TGGCTCAGGT CCCCCTGGGC CCGGAGTGCT  960
CGAGAGCTGT CATGAAGCTG GTCTACTGTG CTCACTGCCT GGGAGTCCCC GGCGCCAGGC 1020
CCTGCCCCGA CTATTGCCGA AATGTGCTCA AGGGCTGCCT TGCCAACCAG GCCAACGTGG 1080
ACGCCGAGTG GAGGAACCTC CTGGACTCCA TGGTGCTCAT CACCGACAAG TTCTGGGGTA 1140
CATCGGGTGT GGAGAGTGTC ATCGGCAGCG TGCACACGTG GCTGGCGGAG GCCATCAACG 1200
CCCTCCAGGA CAACAGGGAC ACGCTCACGG CCAAGGTCAT GGGAACCCCA GGGAACCCCA 1260
AGGTCAACCC CCAGGGCCCT GGGCCTGAGG AGAAGCGGCG CCGGGGCAAG CTGGCCCCGC 1320
GGGAGAGGCC ACCTTCAGGC ACGCTGGAGA AGCTGGTCTC TGAAGCCAAG GCCCAGCTCC 1380
GCGACGTGCC GGACTTCTGG ATCAGCCTCC CAGGGACACT GTGCAGTGAG AAGATGGCCC 1440
TGAGCACTGC CAGTGATGAC CGCTGCTGGA ACGGGATGGC CAGAGGCCGG TACCTCCCCG 1500
AGGTCATGGG TGACGGCCTG GCCAACCAGA TCAACAACCC CGAGGTGGAG GTGGACATCA 1560
CCAAGCCGGA CATGACCATC CGGCAGCAGA TCATGCAGCT GAAGATCATG ACCAACCGGC 1620
TGCGCAGCGC CTACAACGGC AACGACGTGG ACTTCCAGGA CGCCAGTGAC GACGGCAGCG 1680
GCTGGGGCAG CGGTGATGGC TGTCTGGATG ACCTCTGCGG CGGAAGGGTC AGCAGGAAGA 1740
GCTCCAGCTC CCGGACGCCC TTGACCCATG CCCTCCCAGG CCTGTCAGAG CAGGAAGGAC 1800
AGAAGACCTC GGCTGCCAGC TGCCCCCAGC CCCGACCTT CCTCCTGCCC CTCCTCCTCT 1860
TCCTGGCCCT TACAGTAGCC AGGCCCCGGT GGCGGTAACT GCCCCAAGGC CCCAGGGACA 1920
GAGGCCAAGG ACTGACTTTG CCAAAAATAC AACACAGACG ATATTTAATT CACCTCAGCC 1980
TGGAGAGCC TGGGGTGGGA CAGGGAGGGC CGGCGGCTCT GAGCAGGGGC AGGGCGCAGAG 2040
GTCCCAGCCC CAGGCCTGGC CTCGCCTGCC TTTCTGCCTT TTAATTTTGT ATGAGGTCCT 2100
CAGGTCAGCT GGGAGCCAGT GTGCCCAAAA GCCATGTATT TCAGGGACCT CAGGGGCACC 2160
TCCGGCTGCC TAGCCCTCCC CCCAGCTCCC TGCACCGCCG CAGAAGCAGC CCCTCGAGGC 2220
CTACAGAGGA GGCCTCAAAG CAACCCGCTG GAGCCCACAG CGAGCCTGTG CCTTCCTCCC 2280
CGCCTCCTCC CACTGGGACT CCCAGCCAGG CCCACCAGCC AGCCCTGGCC CACCCCCCAG 2340
CCTCCAGGA AGCCCCGCAC GGGCTGTCTG GGTGTCCGCC ATCCAGGGTC TGGCAGAGCC 2400
TCTGAGATGA TGCATGATGC CCTCCCCTCA GCGCAGGCTG CAGAGCCCGG CCCCACCTCC 2460
CTGCGCCCTT GAGGGGCCCC AGCGTCTGCA GGGTGACGCC TGAGACAGCA CCACTGCTGA 2520
GGAGTCTGAG GACTGTCCTC CTACAGACCC TGCAGTGAGG GGCCCTCCAT GCGCAGATGA 2580
GGGGCCACTG ACCCACCTGC GCTTCTGCTG GAGGAGGGGA AGCTGGGCCC AAAGGCCCAG 2640
GGAGGCAGCG TGGGCTCTGC CAATGTGGGC TGCCCCTCGC ACACAGGGCT CACAGGGCAG 2700
GCCTTGCTGG GGTCCAGGGC TGTTGGAGGA CCCCGAGGGC TGAGGAGCAG CCAGGACCCG 2760
CCTGCTCCCA TCCTCACCCA GATCAGGAAC CAGGGGCCTCC CTGTTCACGG TGACACAGGT 2820
CAGGGCTCAG AGTGACCCTC GGCTGTCACC TGCTCACAGG GATGCTGGTG GCTGGTGAGA 2880
CCCCGCACTG CACACGGGAA TGCCTAGGTC CCTTCCCGAC CCAGCCAGCT GCACTGCAGG 2940
GCACGGGGAC CTGGATAGTT AAGGGCTTTT CCAAACATGC ATCCATTTAC TGACACTTCC 3000
TGTCCTTGTT CATGGAGAGC TGTTCGCTCC TCCCAGATGG CTTCGGAGGC CCGCAGGGCC 3060
CACCTTGGAC CCTGGTGACC TCCTGTCACT CACTGAGGCC ATCAGGGCCC TGCCCCAGGC 3120
CTGGACGGCC CCTCCTTCCC TCCTGTGCCC CAGCTGCCAG GTGGCCCTGG GGAGGGGTGG 3180
TGTGGTGTTG GGAAGGGGTC CTGCAGGGGG AGGAGGACTT GGAGGGTCTG GGGGCAGCTG 3240
TCCTGAACCG ACTGACCCTG AGGAGGCCGC TTAGTGCTGC TTTGCTTTTC ATCACCGTCC 3300
CGCACAGTGG ACGGAGGTCC CCGGTTGCTG GTCAGGACCA CATGGCTTGT TCTCTGGAAC 3360
CTGACTTTAG ATGTTTTGGG ATCAGGAGCC CCCAACACAG GCAAGTCCAC CCCATAATAA 3420
CCCTGCCAGT GCCAGGGTGG GCTGGGGGACT CTGGCACAGT GATGCCGGGC GCCAGGACAG 3480
CAGCACTCCC GCTGCACACA GACGGCCTAG GGGTGGCGCT CAGACCCCAC CCTACGGCTCA 3540
TCTCTGGAAG GGGCAGCCCT GAGTGGGTCAC TGGTCAGGGC AGTGGCCAAG CCTGCTGTGT 3600
CCTTCCTCCA CAAGGTCCCC CCACCGCTCA GTGTCAGCGG GTGACGTGTG TTCTTTTGAG 3660
TCCTTGTATG AATAAAAGGC TGGAAACCTA AA
```

Seq ID NO: 248 Protein sequence
Protein Accession #: NP_002072.1

```
        1        11        21        31        41        51
        |         |         |         |         |         |
        |         |         |         |         |         |
MELRARGWWL LCAAAALVAC ARGDPASKSR SCGEVRQIYG AKGFSLSDVP QAEISGEHLR   60
ICPQGYTCCT SEMEENLANR SHAELETALR DSSRVLQAML ATQLRSFDDH FQHLLNDSER  120
TLQATFPGAF GELYTQNARA FRDLYSELRL YYRGANLHLE ETLAEFWARL LERLFKQLHP  180
QLLLPDDYLD CLGKQAEALR PFGEAPRELR LRATRAFVAA RSFVQGLGVA SDVVRKVAQV  240
PLGPECSRAV MKLVYCAHCL GVPGARPCPD YCRNVLKGCL ANQADLDAEW RNLLDSMVLI  300
TDKFWGTSGV ESVIGSVHTW LAEAINALQD NRDTLTAKVI QGCGNPKVNP QGPGPEEKRR  360
```

```
RGKLAPRERP PSGTLEKLVS KAKAQLRDVQ DFWISLPGTL CSEKMALSTA SDDRCWNGMA   420
RGRYLPEVMG DGLANQINNP EVEVDITKPD MTIRQQIMQL KIMTNRLRSA YNGNDVDFQD   480
ASDDGSGSGS
```

5   Seq ID NO: 249 <u>DNA sequence</u>
     Nucleic Acid Accession #: NM_001492.3
     Coding sequence: 8..1864

```
      1         11         21         31         41         51
      |          |          |          |          |          |
10    GAAGGCCATG GTCTCCCCAC GGATGTCCGG GCTCCTCTCC CAGACTGTGA TCCTAGCGCT    60
      CATTTTCCTC CCCCAGACAC GGCCCGCTGG CGTCTTCGAG CTGCAGATCC ACTCTTTCGG   120
      GCCGGGTCCA GGCCCTGGGG CCCCGCGGTC CCCCTGCAGC GCCCGGCTCC CCTGCCGCCT   180
15    CTTCTTCAGA GTCTGCCTGA AGCCTGGGCT CTCAGAGGAG GCCGCCGAGT CCCCGTGCGC   240
      CCTGGGCGCG GCGCTGAGTG CGGCGCGGACC GGTCTACACC GAGCAGCCCG GAGCGCCCGC   300
      GCCTGATCTC CCACTGCCCG ACGGGCTCTT GCAGGTGCCC TTCCGGGACG CCTGGCCTGG   360
      CACCTTCTCT TTCATCATCG AAACCTGGAG AGAGGAGTTA GGAGACCAGA TTGGAGGGCC   420
      CGCCTGGAGC CTGCTGGCGG GCGTGGCTGG CAGGCGGCGC TTGGCAGCCG GAGGCCCGTG   480
20    GGCCCGGGAC ATTCAGCGCG CAGGCGCCTG GGAGCTGCGC TTCTCGTACC GCGCGCGCTG   540
      CGAGCCGCCT GCCGTCGGGA CCGCGTGCAC GCGCCTCTGC CGTCCGCGCA GCGCCCCCTC   600
      GCGGTGCGGT CCGGGACTGC GCCCCTGCGC ACCGCTCGAG GACGAATGTG AGGCGCCGCT   660
      GGTGTGCCGA GCAGGCTGCA GCCCTGAGCA TGGCTTCTGT GAACAGCCCG GTGAATGCCG   720
      ATGCCTAGAG GGCTGGACTG GACCCCTCTG CACGGTCCCT GTCTCCACCA GCAGCTGCCT   780
25    CAGCCCCAGG GGCCGTCCT CTGCTACCAC CGGATGCCTT GTCCCTGGGC CTGGGCCCTG   840
      TGACGGGAAC CCGTGTGCCA ATGGAGGCAG CTGTAGTGAC ACACCCAGGT CCTTTGAATG   900
      CACCTGCCCG CGTGGGTTCT ACGGGCTGCG GTGTGAGGTG AGCGGGGTGA CATGTGCAGA   960
      TGGACCCTGC TTCAACGGCG GCTTGTGTGT CGGGGGTGCG GACCCTGACT CTGCCTACAT  1020
      CTGCCACTGC CCACCTGGTT TCCAAGGCTC CAACTGTGAG AAGAGGGTGG ACCGGTGCAG  1080
30    CCTGCAGCCA TGCCGCAATG GCGGGACTCTG CCTGGACCTG GGCCACGCCC TGCGCTGCCG  1140
      CTGCCGCGCC GGCTTCGCGG GTCCTCGCTG CGAGCACGAC CTGGACGACT GCGCGGGCGG  1200
      CGCCTGCGCT AACGGCGGCA CGTGTGTGGA GGGCGGCGGC GCGCACCGCT GCTCCTGCGC  1260
      GCTGGGCTTC GGCGGGCCGCG ACTGCCGCGA GCGCGCGGAC CCGTGCGCCG CGCGCCCCTG  1320
      TGCTCACGGC GGCCGCTGCT ACGCCCACTT CTCCGGCCTC GTCTGCGCTT GCGCTCCCGG  1380
35    CTACATGGGA GCGCGGTGTG AGTTCCCAGT GCACCCCGAG GGCGCCAAGCG CCTTGCCCGC  1440
      GGCCCCGCCG GGCCTCAGGC CCGGGGACCC TCAGCGCTAC CTTTTGCCTC CGGCTCTGGG  1500
      ACTGCTCGTG GCCGCGGGCG TGGCCGGCGC TGCCGTCTTG CTGGTCCACG TGCGCCGCCG  1560
      TGGCCACTCC CAGGATGCTG GGTCTCGCTT GCTGGCTGGG ACCCGGAGC CGTCAGTCCA  1620
      CGCACTCCCG GATGCACTCA ACAACCTAAG GACGCAGGAG GGTTCCGGGG ATGGTCCGAG  1680
40    CTCGTCCGTA GATTGGAATC GCCCTGAAGA TGTAGACCCT CAAGGGATTT ATGTCATATC  1740
      TGCTCCTTCC ATCTACGCTC GGGAGGTAGC CACGCCCCTT TTCCCCCCGC TACACACTGG  1800
      GCGCGCTGGG CAGGAGGCAGC ACCTGCTTTT TCCCTACCCT TCCTCGATTC TGTCCGTGAA  1860
      ATGAATTGGG TAGAGTCTCT GGAAGGGTTT AAGCCCATTT TCAGTTCTAA CTTACTTTCA
```

45   Seq ID NO: 250 <u>Protein sequence</u>
     Protein Accession #: NP_058637.1

```
      1         11         21         31         41         51
      |          |          |          |          |          |
50    MVSPRMSGLL SQTVILALIF LPQTRPAGVF ELQIHSFGPG PGPGAPRSPC SARLPCRLFF    60
      RVCLKPGLSE EAAESPCALG AALSARGPVY TEQPGAPAPD LPLPDGLLQV PFRDAWPGTF   120
      SFIIETWREE LGDQIGGPAW SLLARVAGRR RLAAGGPWAR DIQRAGAWEL RFSYRARCEP   180
      PAVGTACTRL CRPRSAPSRC GPGLRPCAPL EDECEAPLVC RAGCSPEHGF CEQPGECRCL   240
      EGWTGPLCTV PVSTSSCLSP RGPSSATTGC LVPGPGPCDG NPCANGGSCS ETPRSFECTC   300
55    PRGFYGLRCE VSGVTCADGP CFNGGLCVGG ADPDSAYICH CPPGFQGSNC EKRVDRCSLQ   360
      PCRNGGLCLD LGHALRCRCR AGFAGPRCEH DLDDCAGRAC ANGGTCVEGG GAHRCSCALG   420
      FGGRDCRERA DPCAARPCAH GGRCYAHFSG LVCACAPGYM GARCEFPVHP DGASALPAAP   480
      PGLRPGDPQR YLLPPALGLL VAAGVAGAAL LLVHVRRRGH SQDAGSRLLA GTPEPSVHAL   540
      PDAINNLRTQ EGSGDGPSSS VDWNRPEDVD PQGIYVISAP SIYAREVATP LFPPLETGRA   600
      GQRQHLLFPY PSSILSVK
```

60   Seq ID NO: 251 <u>DNA sequence</u>
     Nucleic Acid Accession #: CAT cluster

```
      1         11         21         31         41         51
      |          |          |          |          |          |
65    GAAATATAAC CATTGCAATT AGAAAATATC CAAAATAGCC TGTTATCTTC CACGTGGCCT    60
      AGATTATTGA CAATCCCAAA TATACAATTT TTCTTTAAAA GTAGTACAAT TTCCTTTGTA   120
      GCTTCAATTC CTTATATGAC TTCAGACTGG AGAAGCCTGT TAAACCACTG TTAGTTTCAG   180
      TTAGAAAGTC TGAGAGACTT TATACATAAA TTCTCAATTT GGCTGCTGTA CACGTGCCAG   240
70    AGTTTTACTA CTGTAGTGAC CGTTGAGAAG ACCCTGTTT ATTTACATTT GAAGCACTGT   300
      TTGTGCAAAC AACCTTTCAT TGTTAAGTGC CTGTATTCCT TTCATTTACT TCATGTCCAG   360
      GGGTGCTATT TACCTAGAAC CATTGTCTAC TACAATTACC ATTTACATTA CAAGTGTGT   420
      GGTTTTCTTT TTCAAGGAGG TTCAATTAAG GCAATAAGAT GTTTGCTGGA GAAACCTATT   480
      GTTTACTGAA AGCACTCAAT GAAGTCAAAT TACTGAAGCT TTTGCCTACA TCTTGGTCTT   540
75    TTATGTAAAT ATGTTAAAAATA TAACATCTAA GGAAAATAAA CAATATTATA ATTATGTGTT   600
      TGCCATTGTC ATATCAAACT TGCTTTGTAT CATACTAATG TTACATAACT TATCGATCAA   660
      TAAAAATACA TTTCAATGTT AAAAAAAAAA AAAAAAA
```

80   Seq ID NO: 252 <u>DNA sequence</u>
     Nucleic Acid Accession #: Eos sequence

```
      1         11         21         31         41         51
      |          |          |          |          |          |
      AGGTACTGCC AGAAAGGATC AGGACCTGGA GTCTGGCAAG AGGAAGACAG AGGCCTGTGT    60
```

```
GGGAAGCGAG TTGTTATCTT TGGTTATCTA GCTGTATGAG TGTATTGGTC TTCATAAAGC   120
TAGATAACCG AAAGTAAAAA CTTCTTCAAG ATCGCCGGGG AGCGTGTGAG AATGAAAGAC   180
TACAGCCGAG AGACAGTAAA AACCAGAAAG GTCAGGAATA CTTATTGAAT CTAACTTTGT   240
TTTTGTTTTG TTTTTTTCCT TATGATTAAA GGTGGGATGA GAGAAAATTA AATGACACAC   300
ACATGCTAAA ATATCAAGGT TCCAGATATG TCTTGAGAGG GGTTGTTGCA GCTGCAAAGG   360
AGAAGTGTAT AGTGATAATG AGTAAAGATG CATGTGCAGT TTGTTCTATT TTAAGGCAAA   420
AGTTATATCA GGGATTTTTT TCTTAGAAAG GTGTTGCAGA GATGTCTGGT ACCTAGTTTA   480
AAAATGATTC CATAATATGT AGACTTGGTC AGTTCCTTTG GGAGGCACCT CCCTCTCAAA   540
ATTTGAAGAT TGTGCTTGGA AATTACTTTA CATGTATTTG GGCTGTATGT CAATTTGTAC   600
GAAATTAGAC TTTCAGAAAA GTTTATACTG GAAGGTTAAT AATTTGTATC TACTGAGGAC   660
TTAGAGCTAG CAGGCAAAAT GAAAAAAAAA AAAAAAAARG CAAGGGCTGA TTTTTATTCT   720
TCTATTCAAA ATACAAGGAC AGATGCTTCT CTGTTCCAAG AGGGTTTCCT TGAGGAAGCT   780
ACTGAAGCAG AAAGACATGA TGGAGACGAG ATCGCCTCCC CTCTTGTCAA AGTGTTAAAA   840
AAATGTTCTG TCTTACTCTG CGCCTAGCAT TGGAAATGAA AGTGACATTT ACGCCACAAC   900
CCACGTGTGC GCCTCCTCTC TTTTTGTTTA AGGATGATCA GGTCTATCCA GGAAACAGCT   960
CTGGCATCCC AAACTGAAAT AATTAGGACG TATATAGACC TGACAAAAAT GGAAAGGGGG  1020
TGGGGAATCT GAGGGTCTGT CTTGCCTAAT TGATTCCGCT AAACGGAATG CAGGAGATGT  1080
GAACGGCAGG ACGCTCCGAT TCCCACGTC GGGGGCAAGT GATAAAGCGG GGCGGGCAG  1140
CCTATGACAG ACAGCCCTGT TGGGGGGTGG GGGTATGAAA AAAACATCAA GTGCACACAC  1200
CATACTCATC TCCATCGCTT AAGAAAGTAA AGGCATTTCC CACCCACAGC CATCTGCAGC  1260
TTCCCAAATG GCAGCACCAA CTGGTCTGTA GCTGCTACAT AGTCTGCTTC TGTTAATTTT  1320
TTAACCACTG TTTAATCTGG CCATAATTAA GTTTGGCTTT CTTCGTTGTT TGAGATTTTC  1380
AGAATTCAAG GCAAGCTAGT AGAAAGCAAT TCCAAGAAAG TCCCATGACT GCCTGCCCCT  1440
AATGTCAAAA TCTCAGTCCA TGAGATTATG GCCTTGTGAC CACATTTTTG CTTTTGTTTT  1500
GGGTGGGCAA ATGTGTATAG AGATAAAATA CATATCTCTA TATAACAGTC GTTATTATAA  1560
TTTCATGAGG CTTTTTACCT CTTAACATGA TACATCTAGG AACTTGGTCT AATTGTGCTA  1620
GTAGATATAC ACTAGAAGTA AAACTATAAA GTCATTAGCC TGTAGAAAGT TGATTATGAT  1680
AACAATATGA TAAAAGTTT GTTTTGGAAT AGTCTCAGCT AGATGGGTCT AAAATAGCCAT  1740
TTTAATGTAA TCTAAAAATA ACACTATGCC TAGCAGAAAC TTTGGCGCTT TGGAGGTCCC  1800
CATTGTCGCC TTTTCATAAA AGTCCTTAAG TTTTCCATAT GTCACCAAGC AAACATTTGT  1860
ATAGGCATTT ACACAGAAAT ATAACCATTG CAATTAGAAA ATATCCAAAA TAGCCTGTTA  1920
TCTTCCACGT GGCCTAGATT ATTGACAATC CCAAATATAC AATTTTTCTT TAAAAGTAGT  1980
ACAATTTCCT TTGTAGCTTC AATTCCTTAT ATGACTTCAG ACTGGAGAAG CCTGTTAAAC  2040
CACTGTTAGT TTCAGTTAGA AAGTCTGAGA GACTTTATAC ATAAATTCTC AATTGGCCTG  2100
CTGTACACGT GCCAGAGTTT TACTACTGTA GTGACCGTTG AGAAGACCCT TGTTTATTTA  2160
CATTTGAAGC ACTGTTTGTG CAAACAACCT TTCATTGTTA AGTGCCTGTA TTCCTTTCAT  2220
TTACTTCATG TCCAGGGGTG CTATTTACCT AGAACCATTG GTTTCATGTT TCTACTACTA  2280
CATTACAAAG TGTGTGGTTT TCTTTTTCAA GGAGGTTCAA TTAAGGCAAT AAGATGTTTG  2340
CTGGAGAAAC CTATTGTTTA CTGAAAGCAC TCAATGAAGT CAAATTACTG AAGCTTTTGC  2400
CTACATCTTG GTCTTTTATG TAAATATGTT AAATATAACA TCTAAGGAAA ATAAACAATA  2460
TTATAATTAT GTGTTTGCCA TTGTCATATC AAACTTGCTT TGTATCATAC TAATGTTACA  2520
TAACTTATCG ATCAATAAAA ATACATTTCA ATGTT
```

Seq ID NO: 253 DNA sequence
Nucleic Acid Accession #: NM_001650.2
Coding sequence; 40.1011

```
     1         11        21        31        41        51
     |         |         |         |         |         |
GGGGCAGGCA ATGGAGAGCTG CACTCTGGCT GGGGAAGGCA TGAGTGACAG ACCCACAGCA   60
AGGCGGTGGG GTAAGTGTGG ACCTTTGTGT ACCAGAGAGA ACATCATGGT GGCTTTCAAA  120
GGGGTCTGGA CTCAAGCTTT CTGGAAAGCA GTCACAGCGG AATTTCTGGC CATGCTTATT  180
TTTGTTCTCC TCAGCCTGGG ATCCACCATC AACTGGGGTG GAACAGAAAA GCCTTTACCG  240
GTCGACATGG TTCTCATCTC CCTTTGCTTT GGACTCAGCA TTGCAACCAT GGTGCAGTGC  300
TTTGGCCATA TCAGCGGTGG CCACATCAAC CCTGCAGTGA CTGTGGCCAT GGTGTGCACC  360
AGGAAGATCA GCATCGGCCAA GTCTGTCTTC TACATCGCAG CCCAGTGCCT GGGGGCCATC  420
ATTGGAGCAG GAATCCTCTA TCTGGTCACA CCTCCCAGTG TGGTGGGAGG CCTGGGAGTC  480
ACCATGGTTC ATGGAAATCT TACCGCTGGT CATGGTCTCC TGGTTGAGTT GATAATCACA  540
TTTCAATTGG TGTTTACTAT CTTTGCCAGC TGTGATTCCA AACGGACTGA TGTCACTGGC  600
TCAATAGCTT TAGCAATTGG ATTTTCTGTT GCAATTGGAC ATTTATTTGC AATCAATTAT  660
ACTGGTGCCA GCATGAATCC CGCCCGATCC TTTGGACCTG CAGTTATCAT GGGAAATTGG  720
GAAAACCATT GGATATATTG GGTTGGGCCC ATCATAGGAG CTGTCCTCGC TGGTGGCCTT  780
TATGAGTATG TCTTCTGTCC AGATGTTGAA TTCAAACGTC GTTTTAAAGA AGCCTTCAGC  840
AAAGCTGCCC AGCAAACAAA AGGAAGCTAC ATGGAGGTGG AGGACAACAG GAGTCAGGTA  900
GAGACGGATG ACCTGATTCT AAAACCTGGA GTGGTGCATA TGATTGACGT TGACCGGGGA  960
GAGGAGAAGA AGGGGAAAGA CCAATCTGGA GAGGTATTGT CTTCAGTATG ACTAGAAGAT 1020
CGCACTGAAA GCAGACAAGA CTCCTTAGAA CTGTCCTCAG ATTTCCTTCC ACCCATTAAG 1080
GAAACAGATT TGTTATAAAT TAGAAATGTG CAGGTTTGTT GTTTCATGTC ATTATTACTA 1140
GTCTAAACAA TAAATATTTC ATAATTTACA AAGGAGGAAC GGAAGAAACC TATTGTGAAT 1200
TCCAAATCTA AAAAAGAAA TATTTTTAAG ATGTTCTTAA GCAAATATAT ACCTATTTTA 1260
TCTAGTTACC TTTCATTAAC AACCAATTTT AACCGTGTGT CAAGATTTGG TTAAGTCTTG 1320
CCTGACAGAA CTCAAAGACA CGTCTATCAG CTTATTCCTT CTCTACTGGA ATATTGGTAT 1380
AGTCAATTCT TATTTGAATA TTTATTCTAT TAAACTGAGT TTAACAATGG C
```

Seq ID NO: 254 Protein sequence
Protein Accession #: NP_001641.1

```
     1         11        21        31        41        51
     |         |         |         |         |         |
MSDRPTARRW GKCGPLCTRE NIMVAFKGVW TQAFWKAVTA EFLAMLIFVL LSLGSTINWG   60
GTEKPLPVDM VLISLCFGLS IATMVQCFGH ISGGHINPAV TVAMVCTRKI SIAKSVFYIA  120
AQCLGAIIGA GILYLVTPPS VVGGLGVTMV HGNLTAGHGL LVELIITFQL VFTIFASCDS  180
KRTDVTGSIA LAIGFSVAIG HLFAINYTGA SMNPARSFGP AVIMGNWENH NIYWVGPIIG  240
AVLAGGLYEY VFCPDVEFKR RFKEAFSKAA QQTKGSYMEV EDNRSQVETD DLILKPGVVH  300
```

VIDVDRGEEK KGKDQSGEVL SSV

Seq ID NO: 255 DNA sequence
Nucleic Acid Accession #: U26742.1
Coding sequence: 325..1449

```
     1         11        21        31        41        51
     |         |         |         |         |         |
     CAGGAAACCC TGGTACTGGC AGCAGCCAGC CTCTGCTGTG CCCACATGAC CCACAACTCT   60
     GGCAGCGGAC CCGGCACTTC CAACATTATT AAATAATAAG AAAGCGGCTC CTACTCCAGG   120
     CTCAAACCTC CCTGCAGACC AATGGACACC TTCTAAGAGT TTGGCGAGTC AGTGACTGAA   180
     GCGCCCGTCC ATTCCAAGAT AAATAGGATT TACCAATCCT TGGATGAAGT GCTTGGGAAG   240
     TCTTTAAGTG CCATAATCAA CTGCCATTTC AAAGAATATA GATGGTTTTG AAAAGTTCAT   300
     GCTGTCCCTT CATTGAATTT TAGAATGATT GAAGATAGTG GGAAAGAGG AAATACCATG   360
     GCAGAAAGAA GACAGCTGTT TGCAGAGATG AGGGCTCAAG ATCTGGATCG CATCCGACTC   420
     TCCACCTACA GAACAGCATG CAAGCTTAGG TTTGTTCAGA AGAAATGCAA TTTGCACCTG   480
     GTGGACATAT GGAATGTCAT AGAAGCATTG CGGGAAAATG CTCTGAACAA CCTGGACCCA   540
     AACACTGAAC TCAACGTGTC CCGCTTAGAG GCTGTGCTCT CCACTATTTT TTACCAGCTC   600
     AACAAACGGA TGCCAACCAC TCACCAAATC CATGTGGAGC AGTCCATCAG CCTCCTCCTT   660
     AACTTCCTGC TTGCAGCGTT TGATCCGGAA GGCCATGGTA AAATTTCAGT ATTTGCTGTC   720
     AAAATGGCTT TAGCCACATT GTGTGGAGGG AAGATCATGG ACAAATTAAG ATATATTTTC   780
     TCAATGATTT CTGACTCCAG TGGGGTGATG GTTTATGGAC GATATGACCA ATTCCTTCGG   840
     GAAGTTCTCA AACTACCCAC GGCAGTTTTT GAAGGTCCTT CATTTGGTTA CACAGAACAG   900
     TCAGCCAGAT CCTGTTTCTC CCAACAGAAA AAAGTCACGT TAAATGGTTT CTTGGACACG   960
     CTTATGTCAG ATCCTCCCCC GCAGTGTCTG GTCTGGTTGC CTCTTCTGCA TCGACTAGCA   1020
     AATGTGGAAA ATGTCTTCCA TCCGGTTGAG TGTTCCTACT GCCACAGTGA GAGTATGATG   1080
     GGATTTCGCT ACCGGATGCCA ACAGTGTCAC AATTACCAGC TCTGTCAGGA CTGCTTCTGG   1140
     AGGGGACATG CCGGTGGTTC TCATAGCAAC CAGCACCAAA TGAAAGAGTA CACGTCATGG   1200
     AAATCACCTG CTAAGAAGCT GACTAATGCA TTAAGCAAGT CCCTGAGCTG TGCTTCCAGC   1260
     CGTGAACCTT TGCACCCCAT GTTCCCAGAT CAGCCTGAGA AGCCACTCAA CTTGGCTCAC   1320
     ATCGTTGATA CTTGGCCTCC CAGACCTGTA ACCAGCATGA ACGACACCCT GTTCTCCCAC   1380
     TCTGTTCCCT CCTCAGGAAG TCCTTTTATT ACCAGGAGCT CGGACGGTGC TTTTGGTGGA   1440
     TGCGTCTAGA TGGATAACAT GACTTCTTCT ACCCTAAAAT ATTCCTATAA TACTTTGAGC   1500
     TGTTCTGGTT CCTCCAGGGT GCATGGTACC CATTAACCCA AAATATGATT ATTTCCCTTT   1560
     TTTCCCATTT TCAGTCATTT TGGAATGTTC TCTGTGAACC ACAGTTGGGT TGTTTAAAGC   1620
     TCACATTTCT TTCTGTCACC ACAGAGATTG GCCTACGGTT TCTGTTTTGA GGGTGCTGTT   1680
     CAATAAAGCT GTGTACACTA AATGTCC
```

Seq ID NO: 256 Protein sequence
Protein Accession #: AAC50424.1

```
     1         11        21        31        41        51
     |         |         |         |         |         |
     MIEDSGKRGN TMAERRQLFA EMRAQDLDRI RLSTYRTACK LRFVQKKCNL HLVDIWNVIE   60
     ALRENALNNL DPNTELNVSR LEAVLSTIFY QLNKRMPTTH QIHVEQSISL LLNFLLAAFD   120
     PEGHGKISVF AVKMALATLC GGKIMDKLRY IFSMISDSSG VMVYGRYDQF LREVLKLPTA   180
     VFEGPSFGYT EQSARSCFSQ QKKVTLNGFL DTLMSDPPPQ CLVWLPLLHR LANVENVFHP   240
     VECSYCHSES MMGFRYRCQQ CHNYQLCQDC FWRGHAGGSH SNQHQMKEYT SWKSPAKKLT   300
     NALSKSLSCA SSREPLHPMF PDQPEKPLNL AHIVDTWPPR PVTSMNDTLF SHSVPSSGSP   360
     FITRSSDGAF GGCV
```

Seq ID NO: 257 DNA sequence
Nucleic Acid Accession #: NM_004172.1
Coding sequence: 179..1807

```
     1         11        21        31        41        51
     |         |         |         |         |         |
     GCGGATTGTT GCTCCGTTGT ACCTGCTGGG GAATTCACCT CGTTACTGCT TGATATCTTC   60
     CACCCCTTAC AAAATCAGAA AAGTTGTGTT TTCTAATACC AAAGAGGAGG TTTGGCTTTC   120
     TGTGGGTGAT TCCCAGACAC TGAAGTGCAA AGAAGAGACC CTCCTAGAAA AGTAAAATAT   180
     GACTAAAAGC AATGGAGAAG AGCCCAAGAT GGGGGGCAGG ATGGAGAGAT TCCAGCAGGG   240
     AGTCCGTAAA CGCACACTTT TGGCCAAGAA GAAAGTGCAG AACATTACAA AGGAGGATGT   300
     TAAAAGTTAC CTGTTTCGGA ATGCTTTTGT GCTGCTCACA GTCACCGCTG TCATTGTGGG   360
     TACAATCCTT GGATTTACCC TCCGACCATA CAGAATGAGC TACCGGGAAG TCAAGTACTT   420
     CTCCTTTCCT GGGGAACTTC TGATGAGGAT GTTACAGATG CTGGTCTTAC CACTTATCAT   480
     CTCCAGTCTT GTCACAGGAA TGGCGGCGCT AGATAGTAAG AGATCAGGGA AGATGGGAAT   540
     GCGAGCTGTA GTCTATTATA TGACTACCAC CATCATTGCT GTGGTGATTG GCATAATCAT   600
     TGTCATCATC ATCCATCCTG GGAAGGGCAC AAAGGAAAAC ATGCACAGAG AAGGCAAAAT   660
     TGTACGAGTG ACAGCTGCAG ATGCCTTCCT GGACTTGATG CGGAACATGT TCCCTCCAAA   720
     TCTGGTAGAA GCCCTGCTTTA AACAGTTTAA AACCAACTAT GAGAAGAGAA GCTTTAAAGT   780
     GCCCATCCAG GCCAACGAAA CGGCTTGTGGG TGCTGTGATA AACAATGTGT CTGAGGCCAT   840
     GGAGACTCTT ACCCGAATCA CAGAGGAGCT GGTCCCAGTT CCAGGATCTG TGAATGGAGT   900
     CAATGCCCTG GGTCTAGTTG TCTTCTCCAT GTGCTTCGGT TTTGTGATTG GAAACATGAA   960
     GGAACAGGGG CAGGCCCTGA GAGAGTTCTT TGATTCTCTT AACGAAGCCA TCATGAGACT   1020
     GGTAGCAGTA ATAATGTGGT ATGCCCCCGT GGGTATTCTC TTCCTGATTG CTGGGAAGAT   1080
     TGTGGAGATG GAAGACATGG TGTGTGATTGG GGGGCAGCTT GCCATGTACA CCGTGACTGT   1140
     CATTGTTGGC TTACTCATTC ACGCAGTCAT CGTCTTGCCA CTCCTCTACT TCTTGGTAAC   1200
     ACGGAAAAAC CCTTGGGTTT TTATTGGAGG GTTGCTGCAA GCACTCATCA CCGCTCTGGG   1260
     GACCTCTTCA AGTTCTGCCA CCCTACCCAT CACCTTCAAG TGCCTGGAAG AGAACAATGG   1320
     CGTGGACAAG CGCGTCACCA GATTCGTGCT CCCCGTAGGA GCCACCATTA ACATGGATGG   1380
     GACTGCCCTC TATGAGGCTT TGGCTGCCAT TTTCATTGCT CAAGTTAACA ACTTTGAACT   1440
     GAACTTCGGA CAAATTATTA CAATCAGCAT CACAGCCACA GCTGCCAGTA TTGGGGCAGC   1500
     TGGAATTCCT CAGGCGGGCC TGGTCACTAT GGTCATTGTG CTGACCATCTG TCGGCCTGCC   1560
     CACTGACGAC ATCACGCTCA TCATCGCTGT GGACTGGTTC CTGGATCGCC TCCGGACCAC   1620
```

```
CACCAACGTA CTGGGAGACT CCCTGGGAGC TGGGATTGTG GAGCACTTGT CACGACATGA    1680
ACTGAAGAAC AGAGATGTTG AAATGGGTAA CTCAGTGATT GAAGAGAATG AAATGAAGAA    1740
ACCATATCAA CTGATTGCAC AGGACAATGA AACTGAGAAA CCCATCGACA GTGAAACCAA    1800
GATGTAGACT AACATAAAGA AACACTTTCT TGAGCACCAG GTGTTAAAAA CCATTATAAA    1860
ATCTTTCCAT CTCATTACAG CTCATTCGCT CCAGCAAGCC CGTCATCTTC CCTTTCCTCC    1920
CTTCTGATAA GACTGGAAAA TAGTCCTCCA AAACACAAGG GAGGATTTTG GGTGGCCAAA    1980
GTGTACAATT TTCATCCCAC AATTGAAATT TTTAAATCAT TTCATGGTTAG TCTTACCGAA    2040
TAAGGTACCA AGATCACAAA TAGTGTTGAT CAGATCTTAC AAGTTTATGT GGCACACAAT    2100
TCCTATAAAT GTGATTTTTT TATATAAGTT AAAGAGACAA ATAGTAGGCT AAAAACATTT    2160
TAAAATCAAC TTTTGAAATT TAAAAATCTT TCAGAATACA ATTCAGTTTT AGTTTCAAAA    2220
TGTTAACAAC TTGAATTACA ACCGGTTATC AGTTGGACAG TAAGATTTTA TCCCTTTCTC    2280
TTCTGACTGG TATACCTATT TCATTAGTAG CTAGGTGCAC ATATACATCT AGCACAGCTG    2340
TGAGGACAGA CAGAAGGCAA AGTTTCCATG TGGCCTTGAG CAAGTCCCAT CTCACCTCTA    2400
GGCCTCAGTG TCCTCATCTA TAAAATGAGG GACTTCCCTA GAAGTCTTCA TGGTCTCTTC    2460
CAGCCCAGAC ATCCTGTGAT GTCATGAAAG CACCTGCCCT CTGTTTCCCC TCAGAACACC    2520
CTGTACCATC CATGGAGCAC GAGGCCTTCA GAAAAGACAC TTCAATGGGA GTGAACATTT    2580
CTAACTAAGG ACAGGATGGC TGTGTGTGGT GGTCACCAGG TCCTGTGAGC AAAGTGCAGG    2640
TTATGCAAGT CGGCCAGGCAG GAGGCCATTC CAGGAGTGGG ATTATTCATC AAACTCTTTG    2700
CCCAGTTCAT CCCAATGGGG GAAGTATTCC CTTCTTTCCT ACTCTGGGAA GAATGTCTCC    2760
TGCCACTCCT CAACTGATGA TAGACTTCGA AAACAGATGA GAAGACTAGC AGCTAGCAAG    2820
GGTGCTTGTA GTCACACTGT GGAACACTAA AGAGCTAGGA AAGAGTTGAG CACAGGCAAC    2880
ATTACAAACA AAGGATTTGA AAACACCAAG AGTACAGGTC TTCTTTAAGG AAGAATAAAA    2940
AAGAAGAGGT TCATTTTTCT GGCTTTTTTT TTCACCTGAA ACACTTTTTC TCGAGTCCAA    3000
AATCATTCCC CCCGTGAAGT CTGCTTACCA AAACATAAGA CGACTTATAT ATTTGAAAGA    3060
AGTCAAATGA ATGAGCTCTC TAATAGAAGT CCATGAGTTG AGTGGGTATT TCTTATTTGA    3120
AAGTGTTTTT CTTTAATCAA AAGTCCTTAG AATGAGGGAA ACAAAATATT TATTTGTTTT    3180
GGAATCCCAC TTATCAAATC ATTCAAAACT TTCAGCTGGA GTGGGGTTTG CTTTTGTTTT    3240
GTTTGTGTCC ATAAGAGAAA TGGTAGAAGA TGAATCAGTA TGAAGACACT GTCAATGAGG    3300
TTATGAGAAA AAAACAGCAG GGGCATTAGT TTCAGGCAAG GCAGCTCCCA GGTTTAGAGA    3360
TTAATTTTTA CCCCCTAAGG AATATCCAGT CAAAGACGCT GAGTGGGAGC TGTCAGGCAG    3420
TAGCAGCTGT GTTTGAGTTT CTGGCTGAAA ATGGTGAAGA ATGGACTTAA TTATGCTAAC    3480
AAACTGAAAA ATCTAGACAT AGATCCTCTG ATATACAATT AGAGATATTT TTATATAGAC    3540
CCCAAGCATT CTGTGCATAA AAGTTAACAT TAGGCTGTGG TGCAGTAACC ATTTAATGTC    3600
GAGGCTCTAT TTCGGAAATA CACTACAAAT GTTAAAGTAC GTGGCTGTCC TCTTAAGACA    3660
CTAGTAGAGC AAAGACTTAA TCATATCAAC TTAATTCTGT TACACAATAT GTGTTTTTTA    3720
ATATACTAAC CATTTCTTAT GGAAAGGTCC TGTGGGGAGC CCATCCTCTC GCCAAGCCAT    3780
CACAGGCTCT GCATACACAT GCACTCAGTG TGGACTGGGA AGCATTACTT TGTAGATGTA    3840
TTTTCAATAA AGAAAAAAAT AGTTTTACAT T
```

Seq ID NO: 258 Protein sequence
Protein Accession #: NP_004163.1

```
    1         11        21        31        41        51
    |         |         |         |         |         |
MTKSNGEEPK MGGRMERFQQ GVRKRTLLAK KKVQNITKED VKSYLFRNAF VLLTVTAVIV    60
GTILGFTLRP YRMSYREVKY FSFPGELLMR MLQMLVLPLI ISSLVTGMAA LDSKASGKMG    120
MRAVVYYMTT TIIAVVIGII IVIIIHPGKG TKEMMEREGK IVRVTAADAF LDLIRNMFPP    180
NLVEACFKQP KTNYEKRSFK VPIQANETLV GAVINNVSEA METLTRITEE LVPVPGSVNG    240
VNALGLVVFS MCFGFVIGNM KEQGQALREF FDSLNEAIMR LVAVIMWYAP VGILFLIAGK    300
IVEMEDMGVI GGQLAMYTVT VIVGLLIHAV IVLPLLYFLV TRKNPWVFIG GLLQALITAL    360
GTSSSSATLP ITFKCLEEMN GVDKRVTRPV LPVGATINMD GTALYEALAA IPIAQVNNFE    420
LNFGQIITIS ITATAASIGA AGIPQAGLVT MVIVLTSVGL PTDDITLIIA VDWFLDRLRT    480
TTNVLGDSLG AGIVEHLSRH ELKNRDVEMG NSVIEENEMK KPYQLIAQDN ETEKPIDSET    540
KM
```

Seq ID NO: 259 DNA sequence
Nucleic Acid Accession #: NM_021948.1
Coding sequence: 48..2783

```
    1         11        21        31        41        51
    |         |         |         |         |         |
TGTGGCACTG CCTGCGTACC CAACCCCAGC CCTGGGTAGC CTGCAGCATG GCCCAGCTGT    60
TCCTGCCCCT GCTGGCAGCC CTGGTCCTGG CCCAGGCTCC TGCAGCTTTA GCAGATGTTC    120
TGGAAGGAGA CAGCTCAGAG GACCGCGCTT TTCGCGTGCG CATCGCGGGC GACGCGCCAC    180
TGCAGGGCGT GCTCGGCGGC GCCCTCACCA TCCCTTGCCA CGTCCACTAC CTGCGGCCAC    240
CGCCGAGCCG CCGGGCTGTG CTGGGCTCTC CGCGGGTCAA GTGGACTTTC CTGTCCCGGG    300
GCCGGGAGGC AGAGGTGCTG GTGGCGCGGG GAGTGCGCGT CAAGGTGAAC GAGGCCTACC    360
GGTTCCGCGT GGCACTGCCT GCGTACCAG CGTCGCTCAC CGACGTCTCC CTGGCGCTGA    420
GCGAGCTGCG CCCCAACGAC TCAGGTATCT ATCGCTGTGA GGTCCAGCAC GGCATCGGATG    480
ACAGCAGCGA CGCTGTGGAG GTCAAGGTCA AAGGGGTCGT CTTTCTCTAC CGAGAGGGCT    540
CTGCCCGCTA TGCTTTCTCC TTTTCTGGGG CCCAGGAGGC CTGTGCCCGC ATTGGAGCCC    600
ACATGCGCAC CCCGGAGCAG CTCTATGCCG CCTACCTTGG GGGCTATGAG CAATGTGATG    660
CTGGCTGGCT GTCGGATCAG ACCGTGAGGT ATCCCATCCA GACCCCACGA GAGGCCTGTT    720
ACGGAGACAT GGATGGCTTC CCCGGGGTCC GGAACTATGG TGTGGTGGAC CCGGATGACC    780
TCTATGATGT GTACTGTTAT GCTGAAGACC TAAATGGAGA ATTGTTCCTG GGTGACCCTC    840
CAGAGAAGCT GACATTGGAG GAAGCACGGG CGTACTGCCA GGAGCGGGGT GCAGAGATTG    900
CCACCACGGG CCAACTGTAT GCAGCCTGGG ATGGTGGCCT GGACCACTGC AGCCAGGGT    960
GGCTAGCTGA TGGCAGTGTG CGCTACCCCA TCGTCACACC CAGCCAGCGC TGTGGTGGGG    1020
GCTTGCCTGG TGTCAAGACT CTCTTCCTCT TCCCCAACCA GACTGGCTTC CCCAATAAGC    1080
ACAGCCGCTT CAACGTCTAC TGCTTCCGAG ACTCGGCCCA GCCTTCTGCC ATCCCTGAGG    1140
CCTCCAACCC AGCCTCCAAC CCAGCCTCTG ATGGACTAGA GGCTATCGTC ACAGTGACAG    1200
AGACCCTGGA GGAACTGCAG CTGCCTCAGG AAGCCACAGA GAGTGAATCC CGTGGGGCCA    1260
TCTACTCCAT CCCCATCATG GAGGACGGAG GAGGTGGAAG CTCCACTCCA GAAGACCCAG    1320
CAGAGGCCCC TAGGACGCTC CTAGAATTTG AAACACAATC CATGGTACCG CCCACGGGGT    1380
```

```
TCTCAGAAGA GGAAGGTAAG GCATTGGAGG AAGAAGAGAA ATATGAAGAT GAAGAAGAGA 1440
AAGAGGAGGA AGAAGAAGAG GAGGAGGTGG AGGATGAGGC TCTGTGGGCA TGGCCCAGCG 1500
AGCTCAGCAG CCCGGGCCCT GAGGCCTCTC TCCCCACTGA GCCAGCAGCC CAGGAGGAGT 1560
CACTCTCCCA GGCGCCAGCA AGGGCAGTCC TGCAGCCTGG TGCATCACCA CTTCCTGATG 1620
GAGAGTCAGA AGCTTCCAGG CCTCCAAGGG TCCATGGACC ACCTACTGAG ACTCTGCCCA 1680
CTCCCAGGGA GAGGAACCTA GCATCCCCAT CACCTTCCAC TCTGGTTGAG GCAAGAGAGA 1740
TGGGGGAGGC AACTGGTGGT CCTGAGCTAT CTGGGGTCCC TCGAGGAGAG AGCGAGGAGA 1800
CAGGAAGCTC CGAGGGTGCC CCTTCCCTGC TTCCAGCCAC ACGGGCCCCT GAGGGTACCA 1860
GGGAGCTGGA GGCCCCCTCT GAAGATAATT CTGGAAGAAC TGCCCCAGCA GGGACCTCAG 1920
TGCAGGCCCA GCCAGTGCTG CCCACTGACA GCGCCAGCCG AGGTGGAGTG GCCGTGGTCC 1980
CCGCATCAGG TGACTGTGTC CCCAGCCCCT GCCACAATGG TGGGACATGC TTGGAGGAGG 2040
AGGAAGGGGT CCGCTGCCTA TGTCTGCCTG GCTATGGGGG GGACCTGTGC GATGTTGGCC 2100
TCCGCTTCTG CAACCCCGGC TGGGACGCCT TCCAGGGCGC CTGCTACAAG CACTTTTCCA 2160
CACGAAGGAG CTGGGAGGAG GCAGAGACCC AGTGCCGGAT GTACGGCTGG CATCTGGCCA 2220
GCATCAGCAC ACCCGAGGAA CAGGACTTCA TCAACAACCG GTACCGGGAG TACCAGTGGA 2280
TCGGACTCAA CGACAGGACC ATCGAAGGCG ACTTCTTGTG GTCGGATGGC GTCCCCCTGC 2340
TCTATGAGAA CTGGAACCCT GGGCAGCCTG ACAGCTACTT CCTGTCTGGA GAGAACTGCG 2400
TGGTCATGGT GTGGCATGAT CAGGGACAAT GGAGTGACGT GCCCTGCAAC TACCACCTGT 2460
CCTACACCTG CAAGATGGGG CTGGTGTCCT GTGGGCCGCC ACCGGAGCTG CCCCTGGCTC 2520
AAGTGTTCGG CCGCCCACGG CTGCGCTATG AGGTGGACAC TGTGCTTCGC TACCGGTGCC 2580
GGGAAGGACT GGCCCAGCGC AATCTGCCGC TGATCCGATG CCAAGAGAAC GGTCGTTGGG 2640
AGGCCCCCCA GATCTCCTGT GTGCCCAGAA GACCTGCCCG AGCTCTGCAC CCAGAGGAGG 2700
ACCCAGAAGG ACGTCAGGGG AGGCTACTGG GACGCTGGAA GGCGCTGTTG ATCCCCCCTT 2760
CCAGCCCCAT GCCAGGTCCC TAGGGGGCAA GGCCTTGAAC ACTGCCGGCC ACAGCACTGC 2820
CCTGTCACCC AAAATTTTCCC TCACACCCTG CGCTCACCAC AGGAAGTGAC AACATGAC
```

Seq ID NO: 260 <u>Protein sequence</u>
Protein Accession #: NP_068767.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MAQLFLPLLA ALVLAQAPAA LADVLEGDSS EDRAFRVRIA GDAPLQGVLG GALTIPCHVH 60
YLRPPPSRRA VLGSPRVKWT FLSRGREAEV LVARGVRVKV NEAYRFRVAL PAYPASLTDV 120
SLALSELRPN DSGIYRCEVQ HGIDDSSDAV EVKVKGVVFL YREGSARYAF SFSGAQRACA 180
RIGAHIATPE QLYAAYLGGY EQCDAGWLSD QTVRYPIQTP REACYGDMDG FPGVRNYGVV 240
DPDDLYDVYC YAEDLNGELF LGDPPEKLTL EEARAYCQEE GAEIATTGQL YAAWDGGLDH 300
CSPGWLADGS VRYPIVTPSQ RCGGGLPGVK TLFLFPNQTG FPNKHSRFNV YCFRDSAQFS 360
AIPEASNPAS NPASDGLEAI VTVTETLEEL QLPQEATESE SRGAIYSIPI MEDGGGGSST 420
PEDPAEAPRT LLEFETQSMV PPTGFSEEEG KALEEEERKYE DEEEKEEEEE EEEVEDEALW 480
AWPSELSSPG PEASLPTEPA AQEESLSQAP ARAVLQPGAS PLPDGESEAS RPPRVHGPPT 540
ETLPTPRERN LASPSPSTLV EAREVGEATG GPELSGVPRG ESEETGSSEG APSLLPATRA 600
PEGTRELEAP SEDNSGRTAP AGTSVQAQFV LPTDSASRGG VAVVPASGDC VPSPCHNGGT 660
CLEEEGVRC LCLPGYYGGDL CDVGLRFCNP GWDAPQGACY KHFSTRRSWE EAETQCRMYG 720
AHLASISTPE EQDFINNRYR EYQWIGLNDR TIEGDFLWSD GVPLLYENWN PGQPDSYFLS 780
GENCVVMVWH DQGQWSDVPC NYHLSYTCKM GLVSCGPPPE LPLAQVFGRP RLRYEVDTVL 840
RYRCREGLAQ RNLPLIRCQE NGRWEAPQIS CVPRRPARAL HPEEDPEGRQ GRLLGRWKAL 900
LIPPSSPMPG P
```

Seq ID NO: 261 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_004386.1
Coding sequence: 2..3967

```
1          11         21         31         41         51
|          |          |          |          |          |
GATGGGGGCC CCGTTTGTCT GGGCCTTGGG CCTTTTGATG CTGCAGATGC TGCTCTTTGT 60
GGCTGGGGAA CAGGGCACAC AGGATATCAC CGATGCCAGC GAAAGGGGGC TCCACATGCA 120
GAAGCTGGGG TCTGGGTCAG TGCAGGCTGC GCTGGCGGAG CTGGTGGCCC TGCCCTGTCT 180
CTTTACCCTG CAGCCACGGC CAAGCGCAGC CCGAGATGCC CCTCGGATAA AGTGGACCAA 240
GGTGCGGACT GCGTCGGCGC AGCGACAGGA CTTGCCCATC CTGGTGGCCA AGGACAATGT 300
CGTGAGGGTG GCCAAAAGCT GGCAGGGACG AGTGTCACTG CCTTCCTACC CCCGGCGCCG 360
AGCCAACGCC ACGCTACTTC TGGGGCCACT GAGGGCCAGT GACTCTGGGC TGTACCGCTG 420
CCAGGTGGTG AGGGGCATCG AGGATGAGCA GGACCTGGTG CCCTTGGAGG TGACAGGTGT 480
TGTGTTCCAC TACCGGATCA GCCCGGGACCG CTATGCACTG ACCTTCGCTG AGGCCCAGGA 540
GGCCTGCCGT CTCAGCTCAG CCATCATTGC AGCCCCTCGG CATCTACAGG CTGCCTTTGA 600
GGATGGCTTT GACAACTGTG ATGCTGGCTG GCTCTCTGAC CGCACTGTTC GGTATCCTAT 660
CACCCAGTCC CGTCCTGGTT GCTATGGCGA CGGTAGCAGC CTTCCAGGGG TTCGGAGCTA 720
TGGGAGGCGC AACCCACAGG AACTCTACGA TGTGTATTGC TTTGCCGGGG AGCTGGGGGG 780
CGAGGTCTTC TACGTGGGCC CGGCCCGCCG CCTGACACTG GCCGGCGCGGC GTGCACAGTG 840
CCGCCGCCAG GGTGCCGCGC TGGCCTCGGT GGGACAGCTG CACCTGGCCT GGCATGGAGGG 900
CCTGGACCAG TGCGACCCGG GCTGGCTGGC CGACGGCAGC GTGCGCTACC CGATCCAGAC 960
GCCGCGGCGG CGCTGCGGGG GCCCCAGCCC GGGCGGTGCGG ACCGTCTACC GCTTCGCTAA 1020
CCGGACCGGC TTCCCCTCAC CCGCCGAGCG CTTCGACGCC TACTGCTTCC GAGCTCATCA 1080
CCCCACGTCA CAACATGGAG ACCTAGAGAC CCCATCCTCT GGGGATGAGG GGGAGATTCT 1140
GTCAGCAGAG GGGCCCCCAG TTAGAGAACT GGAGCCCACC CTGGAGGAGG AAGAGGTGGT 1200
CACCCCTGAC TTCCAGGAGC CTCTGGTGTC CAGTGGGGAA GAAGAAACCC TGATTTTGGA 1260
GGAGAAGCAG GAGTCTCAAC AGACCCTCAG CCCTACCCCT GGGGACCCCA TGCTGGCCTC 1320
ATGGCCCACT GGGGAAGTGT GGCTAAGCAC GGTGGCCCCC AGCCCTAGCG ACATGGGGGC 1380
AGGCACTGCA GCAAGTTCAC ACACGGAGGT GGCCCCAACT GACCCCTATGC CTAGGAGAAG 1440
GGGGCGCTTC AAAGGGTTGA ATGGGCGCTA CTTCCAGCAG CAGGAACGGG AGCCGGGGCT 1500
GCAAGGGGGG ATGGAGGCCA GCGCCCAGCC CCCCACCTCA GAGGCTGCAG TGAACCAAAT 1560
GGAGCCTCCG TTGGCCATGG CAGTCACAGA GATGTTGGGC AGTGGCCAGA GCCGGAGCCC 1620
CTGGGCTGAT CTGACCAATG AGGTGGAATAT GCCTGGAGCT GGTTCTGCTG GTGGCAAGAG 1680
CTCCCCAGAG CCCTGGCTGT GGCCCCCTAC CATGGTCCCA CCCAGCATCC CAGGCCACAG 1740
CAGGGCCCCT GTCCTGGAGC TAGAGAAAGC CGAGGGCCCC AGTGCCAGGC CAGCCACCCC 1800
```

```
        AGACCTGTTT TGGTCCCCCT TGGAGGCCAC TGTCTCAGCT CCCAGCCCTG CCCCCTGGGA 1860
        GGCATTCCCT GTGGCCACCT CCCCAGATCT CCCTATGATG GCCATGCTGC GTGGTCCCAA 1920
        AGAGTGGATG CTACCACACC CCACCCCCAT CTCCACCGAG GCCAATAGAG TTGAGGCACA 1980
        TGGTGAGGCC ACCGCCACGG CTCCACCCTC CCCTGCTGCA GAGACCAAGG TGTATTCCCT 2040
        GCCTCTCTCT TTGACCCCAA CAGGACAGGG TGGAGAGCCC ATGCCCACAA CACCTGAGTC 2100
        CCCCAGGGCA GACTTCAGAG AAACTGGGGA GACCAGCCCT GCTCAGGTCA ACAAAGCTGA 2160
        GCACTCCAGC TCCAGCCCAT GGCCTTCTGT AAACAGGAAT GTGGCTGTAG GTTTTGTCCC 2220
        CACTGAGACT GCCACTGAGC CAACGGGCCT CAGGGGTATC CCGGGGTCTG AGTCTGGGGT 2280
        CTTCGACACA GCAGAAAGCC CCACTTCTGG CTTGCAGGCC ACTGTAGATG AGGTGCAGGA 2340
        CCCCTGGCCC TCAGTGTACA GCAAAGGGCT GGATGCAAGT TCCCCATCTG CCCCCCTGGG 2400
        GAGCCCTGGA GTCTTCTTGG TACCCAAAGT CACCCCAAAT TTGGAGCCTT GGGTTGCTAC 2460
        AGATGAAGGA CCCACTGTGA ATCCCATGGA TTCCACAGTC ACGCCGGCCC CCAGTGATGC 2520
        TAGTGGAATT TGGGAACCTG GATCCCAGGT GTTTGAAGAA GCCGAAAGCA CCACCTTGAG 2580
        CCCTCAGGTG GCCCTGGATA CAAGCATTGT GACGCCCCTC ACGACCCTGG AGCAGGGGGA 2640
        CAAGGTTGGA GTTCCAGCCA TGTCTACACT GGGCTCCTCA AGCTCCCAAC CCCACCCAGA 2700
        GCCAGAGGAT CAGGTGGAGA CCCAGGGAAC ATCAGGAGCT TCAGTGCCTC CGCATCAGAG 2760
        CAGTCCCCTA GGGAAACCGG CTGTTCCTCC TGGGACACCG ACTGCAGCCA GTGTGGGCGA 2820
        GTCTGCCTCA GTTTCCTCAG GGGAGCCTAC GGTACCGTGG GACCCCTCCA GCACCCTGCT 2880
        GCCTGTCACC CTGGGCATAG AGGACTTCGA ACTGGAGGTC CTGGCAGGGA GCCCGGGTGT 2940
        AGAGAGCTTC TGGGAGGAGG TGGCAAGTGG AGAGGAGCCA GCCCTGCCAG GGACCCCTAT 3000
        GAATGCAGGT GCGGAGGAGG TGCACTCAGA TCCCTGTGAG AACAACCCTT GCTTCATGG  3060
        AGGGACATGT AATGCCAATG CACCATGTA TGGCTGTAGC TGTGATCAGG GCTTCGCCGG 3120
        GGAGAACTGT GAGATTGACA TTGATGACTG CCTCTGCAGC CCCTGTGAGA ATGGAGGCAC 3180
        CTGTATTGAT GAGGTCAATG GCTTTGTCTG CCTTTGCCTC CCCAGCTATG GGGGCAGCTT 3240
        TTGTGAGAAA GACACCGAGG CTGTGACCG CGGCTGGCAT AAGTTCCAGG GCCACTGTTA 3300
        CCGGCTATTTT GCCCACCGGA GGGCATGGGA AGATGCCGAG AAGGACTGCC GCCGCCGCTC 3360
        CGGCCACCTG ACCAGCGTCC ACTCACCGGA GGAACACAGC TTCATTAATA GCTTTGGGCA 3420
        TGAAAACACG TGGATCGGCC TGAACGACAG GATCGTCGAG AGAGATTTCC AGTGGACGGA 3480
        CAACACCGGG CTGCAATTTG AGAACTGGCG AGAGAACCAG CCGGACAATT TCTTCGCGGG 3540
        TGGCGAGGAC TGTGTGGTGA TGGTGGCGCA TGAAAGCGGG CGCTGGAACG ATGTCCCCTG 3600
        CAACTACAAC CTACCCTATG TCTGCAAGAA GGGCACAGTG CTCTGTGGTC CCCCTCCGGC 3660
        AGTGGAGAAT GCCTCACTCA TCGGTGCCCG CAAGGCCAAG AACAATGTCC ATGCCACTGT 3720
        AAGGTACCAG TGCAATGAAG GATTTGCCCA GCACCATGTG GTCACCATTC GATGCCGGAG 3780
        CAATGGCAG TGGGACAGGC CCCAAATTGT CTGCACCAAA CCCAGACGTT CACCACCACA 3840
        GCGGGGGACAC CACCACCACC ACCAACACCA CCACCAGCAT CACCACCACA AATCCCGCAA 3900
        GGAGCGCAGA AAACACAAGA AACACCCAAC GGAGGACTGG GAGAAGGACG AAGGGAATTT 3960
        TTGCTGAAGA ACCAGAAAAA AGAAAGCACA ACACCTTTCC CATGCCTCCT CTGGAGCCTT 4020
        CGCCTGGGGA GACAGAACCC AGAGAGAAAC AAGAGAGTCC AGAAGTCCCT GAACCCCAAA 4080
        CTGTTCTCGC AAAAAAAATA TTCCTTTGAA CAAAGGTCTT CTTTTCCTTT TTTTACATAC 4140
        ACAAGATCTT CTTGGCAGGT GGAGCCAGGT GTCTGAAAAG TTCATTCTCG TCTGGCTGAA 4200
        CTCTGGGAGT GTGTCCCAGC TGAGGGAAGC ACAAGTAGCA AAGCTCATTG GTCTGGTCTC 4260
        TTGTTTGCCA GGCTGATTGA AGCAGGCCTT GATGAGGGTG CATGAGTGTA TGTTTGCATT 4320
        CACATGAAGG AATTGCTTTT CACACCAGAA ATTCAGACTT AGTCAATGTT GGCTGAATTC 4380
        CTAAATCCAG GAAGAAGCCT GGACGTAGGG TCATTAGCTT TGGGAATAGA AGGCTACACA 4440
        GAAGCACACT GTTTTTGAAC TTGACAACAG CTCTCCCTTT ACCCTGGACT TCAGCCCAAG 4500
        TTCCGTCTTT GGTCTTGGTG GATAAACACA CAGTGTGGAG ATCCCGACGTA CTGCATTTTA 4560
        GGGATGTTTT TAGGACAACC TCCCTCCATG CCTTCAGAGT TAGGAGTGAG AATGATCAAA 4620
        GCAATATGTA GGTGATGGAG GGAGAGTGTA TTGCTAACCC TTCCAGGTCT AGTCCAGCGC 4680
        TGAGATTTGG TGGTTCTGCA TGTGTGATGA ATCTCTTTCA CACAAATAGA CGAGAGGATA 4740
        TTTAGGGCTA GATGAGCCCA GATTTCTTCC CCCTCCATCT CTCAGGGAGA CAAAGAACCT 4800
        CCTTCCTGGA CCAAGGAGGT GCTGCCAAGT TTTCTAGCCC AGTGCACATA CCCAGTCCTT 4860
        AAGCAGACAT TGGTAGTGCC CCTGCCCTGG GTCCCACTCC TGCCCCACCC CACCCTTGTC 4920
        CCTGGCCATT GCCTGGTGGT CTAGAAACAC TTAAAACTTG AAGTAGTGAC ACCTACCTGC 4980
        GGTCATATTG TAGAGAGATG CTCAGTGTTA AAACTGAAAC ACACAAACAC ACACACACAC 5040
        ACATTTTTCT CTTGTAGATT TTAATTTTTT AAGTGGGAAA GAACTCACCT TGCCTTCCTC 5100
        CCCCAAATGT GCAACCTGTA AAAGGTCTCT CCACACCAGG GGCCAGGATC CAGTTCCCTC 5160
        ATCTCTGGCA GGAAAGATCC ACAGCTTTTC CTCCATGTCT GTTACTCACT TTCAGCAGTC 5220
        CGGGTAAAAT CTGTGGATCA GGGTTAAAAA AGCACCGTGG AGAATGGCCC TCTTCAGGAA 5280
        AGAAAAATAA GCAAATGAAT GGTCCACCTA GGGGTTCAGT AAAGAAAGAA ATGTGTTAAC 5340
        TGAGCCTGAA TCCCTTCTGG GAAGTAATAA TGACCATTGA CAACTAAGAA GTAGACACCA 5400
        TGCTAAAGAC TTACATACAA TCTCCTTGAA TCTTCTCAAT AGCCCATTGA CTTAGAAACT 5460
        GTTACTTTCC CATTTTCACAC ACAGTGAAAC TGAGGCTCAG ATATAAAGGA AAGGTACTGG 5520
        CTTGAAGTCA CAACCACGAC AGGAGTAAGG ATTTGGAATA AGGATTTGGT CCTGTTTTCT 5580
        GGACCAAATC CTTACTCTGG CTCTGCTTAC ACTTTCTCTC CATCACCAAA TCCTTACTCC 5640
        AAATCCAGAA GTCAGAGCCA ACTCCCATCT TGGTTCTGAC CCAAATCCTG CTCTGGACTC 5700
        TGGAGAGGAG ATTGAAATAT AATTGCACCC TCATACACAT TTAGGAAATG GTTAAGAAGT 5760
        GTAAACTGAA CCCTTATCCT TGTCTTCAAT CTTCCTCCCT GTAGACATCT ATCTTATTAT 5820
        GGTTATTATT CAGAAAACCC AGGGATACAG GTTTGTCTTC TTACTTTGAT AACTCTTCTT 5880
        AGTTTAAAAT AATAATAATA ACACATCTTT GGTCATCTAT GTCACACAAA AATTTTCCTT 5940
        TGTTTGCGGG GGGCTGGGGA TGCAGTGTTT TTTGGGGGGT CTTGGTTTAT GCTCCCTGCC 6000
        CTTGAGCCCC TCAGCCGTTT GCCCTGCCCC CACCTCGGCT CCATGGTGGG AGGGGGCTCT 6060
        GGTCTTTTCT AAAGTGGGCG GTTTGTCTTT TGATCTTTCC CTTTTGGATG TGCGTGTGTG 6120
        TCTGCGTGTG CCATGTGCGT GGCACGCATA TGAGTGTGTG TGCGTGTGAA CGGCTTTGGG 6180
        TCCTGCTGGT TTTGCTGTGA GCTGCAGTGT TCTGTGGGTC TGTGGTATCT GACACTGTGG 6240
        ACATTAATGT ACTTCTTGGA CATTTTAATA AATTTTTTAA CAGTTCAAAA AAAAAAAAAA 6300
        AAAAAAAAAA
```

Seq ID NO: 262 Protein sequence
Protein Accession #: NP_004377.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MGAPFVWALG LLMLQMLLFV AGEQGTQDIT DASERGLHMQ KLGSGSVQAA LAELVALPCL    60
```

```
      FTLQPRPSAA  RDAPRIKWTK  VRTASGQRQD  LPILVAKDNV  VRVAKSWQGR  VSLPSYPRRR   120
      ANATLLLGPL  RASDSGLYRC  QVVRGIEDEQ  DLVPLEVTGV  VFHYRSARDR  YALTFAEAQE   180
 5    ACRLSSAIIA  APRHLQAAFE  DGFDNCDAGW  LSDRTVRYPI  TQSRPGCYGD  RSSLFGVRSY   240
      GRRNPQELYD  VYCFARELGG  EVPYVGPARR  LTLAGARAQC  RRQGAALASV  GQLHLAWHEG   300
      LDQCDPGWLA  DGSVRYPIQT  PRRRCGGPAP  GVRTVYRFAN  RTGFPSPAER  FDAYCFRAIH   360
      PTSQHGDLET  PSSGDEGEIL  SAEGPPVREL  EPTLEREEVV  TPDFQEPLVS  SGKEETLILE   420
10    EKQESQQTLS  PTPGDPMLAS  WPTGEVWLST  VAPSPSDMGA  GTAASSHTEV  APTDPMPRRR   480
      GRFKGLNGRY  FQQQEPEPGL  QGGMEASAQP  PTSEAAVNQM  EPPLAMAVTE  MLGSGQSRSP   540
      WADLTNEVDM  PGAGSAGGKS  SPEPWLWPPT  MVPPSISGHS  RAPVLELEKA  EGPSARPATP   600
      DLFWSPLEAT  VSAPSPAPWE  AFPVATSPDL  PMMAMLRGPK  EWMLPHPTPI  STEANRVEAH   660
      GEATATAPPS  PAAETKVYSL  PLSLTPTGQG  GEAMPTTPES  PRADFRETGE  TSPAQVNKAE   720
      HSSSSPWPSV  NRNVAVGFVP  TETATEPTGL  RGIPGSESGV  FDTAESPTSG  LQATVDEVQD   780
15    PWPSVYSKGL  DASSPSAPLG  SPGVFLVPKV  TPNLEPWVAT  DEGPTVNPMD  STVTPAPSDA   840
      SGIWEPGSQV  FEEAESTTLS  PQVALDTSIV  TPLTTLEQED  KVGVPAMSTL  GSSSSQPHPS   900
      PEDQVETQGT  SGASVPPHQS  SPLGKPAVPP  GTPTAASVGE  SASVSSGEPT  VPWDPSSTLL   960
      PVTLGIEDFE  LEVLAGSPGV  ESFWEEVASG  EEPALPGTPM  NAGAEEVHSD  PCENNPCLHG  1020
      GTCNANGTMY  GCSCDQGFAG  ENCEIDIDDC  LCSPCENGGT  CIDEVNGFVC  LCLPSYGGSF  1080
20    CEKDTEGCDR  GWHKFQGHCY  RYFAHRRAWE  DAEKDCRRRS  GHLTSVHSPE  EESFINSFGH  1140
      ENTWIGLNDR  IVERDFQWTD  NTGLQFENWR  ENQPDNPFAG  GEDCVVMVAH  ESGRWNDVPC  1200
      NYNLPYVCKK  GTVLCGPPPA  VENASLIGAR  KAKNNVHATV  RYQCNEGFAQ  HHVVTIRCRS  1260
      NGKWDRPQIV  CTKPRRSHRM  RGHHHHHQHH  HQHHHHKSRK  ERRKHKKHPT  EDWEKDEGNF  1320
      C
```

Seq ID NO: 263 <u>DNA sequence</u>
Nucleic Acid Accession #: Eos sequence
Coding sequence: 7..2085

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      GCCGCGATGG  CCAGCACCAG  GAGTATCGAG  CTGGAGCACT  TTGAGGAACG  GGACAAAAGG    60
      CCGCGGCCGG  GGTCGCGGAG  AGGGGCCCCC  AGCTCCTCGG  GGGGCAGCAG  CAGCTCGGGC   120
      CCCAAGGGGA  ACGGGCTCAT  CCCCAGTCCG  GCGCACAGTG  CCCACTGCAG  CTTCTACCGC   180
      ACGCGGACCC  TGCAGGCCCT  CAGCTCGGAG  AAGAAGGCCA  AGAAGGCGCG  CTTCTACCGG   240
      AACGGGGACC  GCTACTTCAA  GGGCCTGGTG  TTTGCCATCT  CCAGCGACCG  CTTCCGGTCC   300
35    TTCGATGCGC  TCCTCATAGA  GCTCACCCGC  TCCCTGTCGG  ACAACGTGAA  CCTGCCCCAG   360
      GGTGTCCGCA  CTATCTACAC  CATCGACGGC  AGCCGGAAGG  TCACCAGCCT  GGACGAGCTG   420
      CTGGAAGGTG  AGAGTTACGT  GTGTGCATCC  AATGAACCAT  TTCGTAAAGT  CGATTACACC   480
      AAAAATATTA  ATCCAAACTG  GTCTGTGAAC  ATCAAGGGTG  GGACATCCCG  AGCGGCTGGCT   540
      GCTGCCTCCT  CTGTGAAAAG  TGAAGTAAAA  GAAAGTAAAG  ATTTCATCAA  ACCCAAGTTA   600
40    GTGACTGTGA  TTCGAAGTGG  AGTGAAGCCT  AGAAAAGCCG  TGCGGATCCT  TCTGAATAAA   660
      AAGACTGCTC  ATTCCTTTGA  ACAAGTCTTA  ACAGATATCA  CCGAAGCCAT  TAAACNAGCC   720
      TCAGGAGTCG  TCAAGAGGCT  CTGCACCCTG  GATGGAAAGC  AGGTGAGAGT  TACGTGTGTG   780
      CATCTGCCAG  ACTTTTTTGG  TGATGACGAT  GTTTTTATTG  CATGTGGACC  AGAAAAATTT   840
      CGTTATGCCC  AAGATGACTT  TGTCCTGGAT  CATAGTGAAT  GTCGTGTCCT  GAAGTCATCT   900
45    TATTCTCGAT  CCTCAGCTGT  TAAGTATTCT  GGATCCAAAA  GCCCTGGGCC  CTCTCGACGC   960
      AGCCAGATTT  CTGCTCATGG  CAGATCTTCT  TCCAATGTAA  ACGGTGGACC  TGAGCTTGAC  1020
      CGTTGCATAA  GTCCTGAAGG  TGTGAATGGA  AACAGATGCT  CTGAATCATC  AACTCTTCTT  1080
      GAGAAATACA  AAATTGGAAA  GGTCATTGGT  GATGGCAATT  TTGCAGTAGT  CAAAGAGTGT  1140
      ATAGACAGGT  CCACTGGAAA  GGAGTTTGCC  CTAAAGATTA  TAGACAAAGC  CAAATGTTGT  1200
50    GGAAAGGAAC  ACCTGATTGA  GAATGAAGTG  TCAATACTGC  GCCGAGTGAA  ACATCCCAAT  1260
      ATCATTATGC  TGGTCGAGGA  GATGGAAACA  GCAACTGAGC  TCTTTCTGGT  GATGGAATTG  1320
      GTCAAAGGTG  GAGATCTCTT  TGATGCAATT  ACTTCGTCGA  CCAAGTACAC  TGAGAGAGAT  1380
      GGCAGTGCAC  TGGTGTACAA  CTTAGCCAAT  GCCCTCAGGT  ATCTCCATGG  CCTCAGCATC  1440
      GTGCACAGAG  ACATCAAACC  AGAGAATCTC  TTGGTGTGTG  AATATCCTGA  TGGAACCAAG  1500
55    TCTTTGAAAC  TGGGAGACTT  TGGGCTTGCG  ACTGTGGTAG  AAGGCCCTTT  ATACACAGTC  1560
      TGTGGCACAC  CCACTTATGT  GGCTCCAAAA  ATCATTGCTG  AAACTGGCTA  TGGCCTGAAG  1620
      GTGGACATTT  GGGCAGCTGG  TGTGATCACA  TACATACTTC  TCTGTGGATT  CCCACCATTC  1680
      CGAAGTGAGA  ACAATCTCCA  GGAAGATCTC  TTCGACCAGA  TCTTGGCTGG  GAAGCTGGAG  1740
      TTTCCGGCCC  CCTACTGGGA  TAACATCACG  GACTCTGCCA  AGGAATTAAT  CAGTCAAATG  1800
60    CTTCAGGGTAA  ATGTTGAAGC  TGGGTGTACC  GCGGGGACAAA  TCCTGAGTCA  CCCCTGGGTG  1860
      TCAGATGATG  CCTCCCAGGA  GAATAACATG  CAAGCTGAGG  TGACAGGTAA  ACTAAAACAG  1920
      CACTTTAATA  ATGCGCTCCC  CAAACAGAAC  AGCACTACCA  CCGGGGTCTC  CGTCATCATG  1980
      GTGAGTGGAA  GGCGGCAGGT  CTGGCCTGAC  TGCGGAGCCG  GCCTTGAAGT  TTTTGAATTA  2040
      GGTAGCCGGG  AGCTGCCCTC  ACATGGAAGT  TGGTGCCTTC  CGTAGTCCTA  TTTCATATGA  2100
65    AGATTGGCTT  GGCATGTGGA  GGGCACTCAT  TCGGCAACTC  CCAGGCTTTG  GGCACTGTGT  2160
      GGAGGGGCTT  GTGTAGGGAC  CAGCAGGCCT  GGTGTGAGGG  GTCCAGGCGT  CAAGGAGCTC  2220
      CTGGCTGGGC  CCTCTGGGCA  GCTGCTTCCA  CTCTTGTCTC  TGCCTTCTCA  TCTAGAGAGA  2280
      CTCCCAAGCC  CTGGAGGGGT  GTGTTGTGTT  AGGAATTAAC  TCCCTGCCTA  CCCCAAGGCC  2340
      TCAGAAATAG  ATTATTAGTT  ATGTGAATTA  TTCTTTGAGA  CTTGGGATAA  GAAACAGCCA  2400
70    AAGCTAAACA  TATTTCAGTT  TTAAAAAATC  AGTGTTTTAT  AAAACACAGT  TTGGGGCTTA  2460
      TAAAGGTACA  TAATCAAGGA  AAAAAATATA  TATTCATTTT  TCAGGGTTGG  TAACATTTTA  2520
      TGAGATGTCA  GTGACAACGA  TGGCCTTATT  TTTTTCAGGG  TTTTCTTCTT  CCAAAATGTT  2580
      TCTTAAGGCA  ACTCTCCTAA  ATACATAAAC  ACAACAAATT  AAAATGAAAA  GTGACATGAG  2640
      AGTAAATGAA  TCAAAAGGAA  AAAACATTGA  ACCAGAGGTG  AGGGCAGCAC  ACCCGCAGCA  2700
75    GCTGTCCAGG  CCTGAGCCAA  TGCAACCCTG  GGCGGGAAGG  CCAGCTCACC  GTGAGCAGGT  2760
      AGAAGCCAGC  CAGCCACCCA  GGCAGGGACC  TTGGTTCTCC  CCACACACTC  CCAGGAGCAG  2820
      GGAACAGGGG  TGGAGTGGCC  TTTCCCAGAG  CTGGAGTTGG  CTGCAGCAGC  TTTCGAATCA  2880
      GACCTGCCAA  GGTGATGGGC  GTCGAGTTT  CACACTCGGG  CCCCCGTGA  CCCCACTGAG  2940
      TCCTGACAGC  TAAGGATGGG  CCACCTCCAC  AGCTCCGTCA  CTCGTACTTG  GGACAGGCCT  3000
80    CTCATCCTCT  GGGAAGGTCC  TCCTTGTTTC  CTACCCAACT  AGAAGGGAAA  CAGTGGCATA  3060
      TTCTCATGGT  ACATGGTTGT  CTGAAAGCCT  TACCTAGGAA  GACGCAGGGT  CTAGATAGAA  3120
      GCTATAAGGA  AGCCACACAC  ATAACCCACA  TCCCCACACC  CCCAACATCC  CCCACACTCC  3180
      CCACACCCCC  CACACCCCCC  ACATCCCCAC  CATAATTACC  CCCACCTCCA  AATATCTCAT
```

Seq ID NO: 264 Protein sequence
Protein Accession #: Eos sequence

```
      1          11         21         31         41         51
      |          |          |          |          |          |
 5    MASTRSIELE HFEERDKRPR PGSRRGAPSS SGGSSSSGPK GNGLIPSPAH SAHCSFYRTR   60
      TLQALSSEKK AKKARFYRNG DRYFKGLVPA ISSDRPRSFD ALLIELTRSL SDNVNLPQGV  120
      RTIYTIDGSR KVTSLDELLE GESYVCASNE PFRKVDYTKN INPNWSVNIK GGTSRALAAA  180
      SSVKSEVKES KDFIKPKLVT VIRSGVKPRK AVRILLNKKT AHSFEQVLTD ITEAIKKASG  240
10    VVKRLCTLDG KQVRVTCVHL PDFFGDDDVF IACGPEKFRY AQDDFVLDHS ECRVLKSSYS  300
      RSSAVKYSGS KSPGPSRRSQ ISAEGRSSSN VNGGPELDRC ISPEGVNGNR CSESSTLLEK  360
      YKIGKVIGDG NFAVVKECID RSTGKEFALK IIDKAKCCGK EHLIENEVSI LRRVKHPNII  420
      MLVEEMETAT ELFLVMELVK GGDLFDAITS STKYTERDGS AMVYNLANAL RYLHGLSIVH  480
      RDIKPENLLV CEYPDGTKSL KLGDFGLATV VEGPLYTVCG TPTYVAPXII AETGYGLKVD  540
15    IWAAGVITYI LLCGFPPFRS ENNLQEDLFD QILAGKLEFP APYWDNITDS AKELISQMLQ  600
      VNVEARCTAG QILSHPWVSD DASQENNMQA EVTGKLKQHP NNALPKQNST TTGVSVIMVS  660
      GRRQVWPDCG AGLEVFELGS RELPSHGSWC LP
```

Seq ID NO: 265 DNA sequence
20 Nucleic Acid Accession #: AB020684.1
Coding sequence: 1..1744

```
      1          11         21         31         41         51
      |          |          |          |          |          |
25    CCCCCTTGTC ATTAATACAT TAAAAAGATT CAATCTTTAC CCTGAGGTAA TTTTGGCCAG   60
      TTGGTACCGG ATTTATACCA AAATAATGGA CTTGATTGGT ATTCAAACCA AGATATGTTG  120
      GACGGTTACC AGAGGAGAAG GACTCAGTCC TATTGAAAGC TGTGAAGGAT TGGGAGATCC  180
      TGCTTGCTTT TATGTTGCTG TAATTTTTAT TTTAAATGAT CTAATGATGG CATTATTCTT  240
      CATATATGGC ACATATTTAA GTGGCAGCCG ATTAGGAGGC CTGGTTACAG TGTTGTGCTT  300
30    CTTTTTCAAT CATGGAGAGT GTACCCGTGT AATGTGGACA CCACCTCTCC GTGAAAGCTT  360
      CTCATATCCA TTTCTTGTTC TTCAGATGTT GGCTAGTGACT CATATTCTCA GGGCTACAAA  420
      ACTTTATAGA GGAAGCTTGA TTGCACTCTG CATTTCCAAT GTATTTTTCA TGCTTCCTTG  480
      GCAGTTTGCT CAGTTTGTAC TTCTTACTCA GATTGCATCA TTATTGCAG TATATGTTGT  540
      CGGGTACATT GATATATGTA AATTACGGAA GATCATTTAT ACACACATGA TTTCTCTTGC  600
35    ACTTTGTTTT GTTTTGATGT TTGGGAACTC AATGTTATTA ACTTCTTATT ATGCTTCTTC  660
      TTTGGTAATT ATTTGGGGTA TTCTGGCAAT GAAACCACAT TTCCTGAAAA TAAATGTATC  720
      TGAACTTAGT TTATGGGTTA TTCAAGGATG TTTTTGGTTA TTTGGAACTG TCATACTTAA  780
      ATACTTGACA TCTAAAATTT TTGGTATTGC AGATGACGCT CATATTGGCA ACTTACTAAC  840
      ATCAAAATTC TTTAGTTATA AGGATTTTGA TACTTTATTG TATACCTGTG CAGCGGAGTT  900
40    TGACTTTATG GAAAAAGAGA CTCCCACTGAG ATACACAGG ACATTATTGC TTCCAGTTGT  960
      TCTTGTAGTG TTTGTTGCTA TTGTTAGAAA GATTATTAGT GATATGTGGG GTGTCTTAGC 1020
      TAAACAACAG ACACATGTAA GAAAACACCA GTTTGATCAT GGAGAGCTGG TTTACCATGC 1080
      ATTGCAATTG TTAGCATATA CAGCCCTTGG TATTTTAATT ATGAGACTAA AACTCTTCTT 1140
      GACACCACAC ATGTGTGTTA TGGCATCACT GATCTGCTCA AGACAGCTAT TTGGATGGCT 1200
45    CTTTTTGCAAA GTACATCCTG GTGCTATTGT GTTTGCTATA TTAGCAGCAA TGTCAATACA 1260
      AGGTTCAGCA AATCTGCAAA CCCAGTGGAA TATTGTAGGG GAGTTCAGCA ATTTGCCCCA 1320
      AGAAGAACTT ATAGAATGGA TCAAATATAG TACTAAACCA GATGCAGTGT TTGCGGGTGC 1380
      CATGCCCACG ATGGCAAGTG TTAAGCTCTC TGCCACTTCGG CCCATTGTGA ATCATCCACA 1440
      TTATGAAGAC GCAGGCTTGA GAGCCAGAAC AAAAAATAGTA TACTCAATGT ATAGTCGGAA 1500
50    AGCAGCCGAA GAAGTGAAGC GAGAACTGAT AAAGTTAAAA GTGAACTATT ACATTCTAGA 1560
      AGAGTCATGG TGTGTAAGAA GATCCAAGCC TGGTTGCAGT ATGCCTGAAA TTTGGGATGT 1620
      AGAAGATCCT GCCAATGCTG GGAAAACTCC CTTATGTAAC CTCTTGGTGA AGGATTCCAA 1680
      ACCTCACTTC ACCACTGTAT TCCAGAACAG TGTTTAAACAA GTCCTAGAAG TTGTAAAAGA 1740
      ATGACTGCTA CATGACCTGC TGCCTACGGA GAACTACATC TGTAATGGTT TTAATGTTTT 1800
55    GCTAAGTCAT GTGTTGTTCA TATCCCAAAA ACTTTTTATG GTAACTGTTT TCAAATAGAA 1860
      AACGTTTTAT TTGGTCAATT TGAATGTCAT TCTAATTATA AAAATGACTT ACACCTTTAT 1920
      CAATTGGTTA CTATTTCAAT GCACCCTTTA AAATTTGCTA TGCAAATGAG TATATGCTTG 1980
      TACTTGACTT TAATATTTGT GCTAAAGTGA GCAAAGCTAC CTGTATAAAG AAAACACAGT 2040
      GGGTTGTGAC AAGGATGACA TGAAAATACA GGACAATTCT GACAATGTAG GGGCTGATTT 2100
60    TATAGTGTAA GAACTATTAA TGCCCCTTGC TTCTTTTTTC TGCCTCTTGC TCTTGTCTTT 2160
      TGGACATTTC AGTGATTGTA AGTTCTTCGG TCATGTCAGC CCCTGTCATC AACTTGAGTT 2220
      ACAGTAGATG GGGCAGACAT GGAGTGTTTG CTATATAAAA CTATCTGTTT GTTTTACTTG 2280
      CTTGTCGCGCT TTTTGTTCTC TGTTCTCTTG TTAATGAAGC TTTTCCTGCC CATTATTAAT 2340
      CCAAACTCTT GGACCTTGTG GTTAGGAAAT TCCCTTAACT TCCAGCCATA TGGCATTATC 2400
65    GTGTCTCTTT CTCTCTCTCT CTGCCTCTCT CTCTTCTCGT CTTCCCCATA TTTTCTGTCA 2460
      AATAAGTACT GTTTACTCAT TTAGTTGCTT ATCAAGTACT TATTCTTGGT TTTAAAAAAA 2520
      ATTAATGGTA ACTGTATTTT TCTCATTTTT AGCATTATTC AAATGTTTAT ATTTTAATAC 2580
      CTTTAAACCA CTTTAAAGTT TTTTCATGTT TAATTATAGT TTTAAGAAAA ACTATTTTGA 2640
      ACAACCCCAA ATATAGTAGTA TCTAGAAACT AATGTATATT TGATTAGACA TCATTTAAGT 2700
70    TGGAACAGTA GACTGTAGTA CATGGTAATT TTTCTTTTAC TATTAAGATA CAATAAAACA 2760
      TGACTAATTT TGCTGTCAAA AATGTAAAGA ATAATGATAA ATGGAGTTTT TTATATTTTA 2820
      CTTTTAAGAT TGCCTGTCTT TAATAAGACA AAGCCTTAAG CCTTATGTTA TAATTTTGGT 2880
      TCTAAAAACC ATCATTTCAG TATAAGGAAT AAGTATATTT CGTCCTCCTC TTTAGTTTTT 2940
      TTCTTCCTAT TTATTTTTAT TTTGAAAAAT TTCTACACCT TCTTTGAATT CCTTGTATGA 3000
75    ATTTTTGTTT CTAGAAGTT AATTTGTGTG AAATGAGATT CTTCAAACG ATGAAACCTC 3060
      ATAGCTCTGA GAAAAGGTTT TAGGGTTTTA AATTCTAAGC AAAGCGTGAC TATGGCTGAC 3120
      AGACTACACA TTTAATTATA CAGCTTCTCT TTCTTAACCA CAGGCAGATT AACCTCATTG 3180
      TGGATTGTCC TTCAGACCTT AGTCCTCAGG CATGGTTTCT CCTGCCCACT CCTGGAAGCC 3240
      GCTGTTCCCT TTCTACCTTC TTACCAGAGC CCAAGGGCAG GCCTGGTCCC GGGGAAGCCA 3300
80    CAGCTTGCTG ACATAAGTCA GCTGCAAAGG CTGAGGAGTG TGCCCTCAGA GAAGCACCGC 3360
      CCCCCAGTCT TGTGCCAGCG CCTAGAGCCG CAGCTCCCAG GGATGCTCCT TCCCTGGAAG 3420
      CAGCCCAGGA GAGGGACTCT GGCAGCGGTTC TTCAGATTTG TGGCCACTGT TTCTCATTTG 3480
      CTGGTTGACT GTTTTTATTT CTTAGGCTTT TGCTAGTTTT AGAAAATAGG GAAGCAGCCC 3540
      TTGATTTGTG GATTAAAAGC AACATTTGAG CGATGATGCA CAACAGTCCA GGAAAATGGG 3600
```

```
CGGTGGACAC TTGAGGCTGA GGATGGGAGT TGACATGAGC AGGGAGAGGG AGGTGCGCGC   3660
TGCTTATCTG TGATTGTTGC TCACCTGAGT GTGGCTGATT GTGTACATCC AGCAGTTACA   3720
ATTTTTAAAA ATTATACTTT TACATTTATT TTATATTTTT CTCACCCCCA GTAATTTCCT   3780
TCCAAAGAAG TTCACATGTA ATAAGTAGAA ATTCTGTATA GGAAAAAAGC ATTAAAAATA   3840
CTATTATAAC TGCTTCATTT GCTGGGAACC ATTAAAAGTA ATATAAATTA GCTTTTTCCA   3900
GAAGGATCCT TTTGTAGCAG TGTTTATGAA TGTAACCCCC AGCAAAATAT GGCTATATAT   3960
TAGGGGAGCC AGTTTGGAGC AGAGGCCTGA AGGTCCCTGC TATGCAGCCG TGGCCACAGC   4020
TCGCAGCCCCA AGCACTGTGG AGCATCCCAA CCTTTGATGG CAATGCAGAT TGGTAGCAGG   4080
TTCCATAGGC GTACAAAACA GTATTAAAGC TCAGTGTTTT GCATATTGTT AGCATTTACA   4140
AATATTTTTG CTTTAGTATG AGGAAAGTAA GGATGGGCAA AGAAGCGATC AAAATAGCTA   4200
TTGCTACAAC ATTTTCGAAA ACAAAGTTGG GGCTGTTATT CTTTAAAAAG ATAAGCCTCT   4260
AAAAATGCTT GGCAAAAAAA ATATAGTGTT AAAATAGGCC AGTGATATTA ATGAGAAAAT   4320
GAAAGTATGT ATCAGGAATA AAGTGATATT GCATAGGAGT ATTGTATTTT TATGAATTTT   4380
ATGCCAGTTG TTTACATGTA CTATATATGT TAAATTAAAA AAAATCATGA GAAATG
```

Seq ID NO: 266 Protein sequence
Protein Accession #: BAA74900.1

```
1          11         21         31         41         51
|          |          |          |          |          |
PLVINTLKRF NLYPEVILAS WYRIYTKIMD LIGIQTKICW TVTRGEGLSP IESCEGLGDP    60
ACFYVAVIFI LNGLMMALPP IYGTYLSGSR LGGLVTVLCF FFNHGECTRV MWTPPLRESF   120
SYPFLVLQML LVTHILRATK LYRGSLIALC ISNVFFMLPW QFAQFVLLTQ IASLFAVYVV   180
GYIDICKLRK IIYIHMISLA LCFVLMFGNS MLLTSYYASS LVIIWGILAM KPHPLKINVS   240
ELSLWVIQGC FWLFGTVILK YLTSKIFGIA DDAHIGNLLT SKFFSYKDFD TLLYTCAAEF   300
DFMEKETPLR YTKTLLLPVV LVVFVAIVRK IISDMWGVLA KQQTHVRKHQ FDHGELVYHA   360
LQLLAYTALG ILIMRLKLFL TPHMCVMASL ICSRQLFGWL FCKVHPGAIV FAILAAMSIQ   420
GSANLQTQWN IVGEFSNLPQ EELIBWIKYS TKPDAVFAGA MPTMASVKLS ALRPIVNHPH   480
YEDAGLRART XIVYSMYSRK AAEEVKRELI KLKVNYYILE ESWCVRRSKP GCSMPEIWDV   540
EDPANAGKTP LCNLLVKDSK PHFTVVFQNS VYKVLEVVKE
```

Seq ID NO: 267 DNA sequence
Nucleic Acid Accession #: U26744.1
Coding sequence: 59..1600

```
1          11         21         31         41         51
|          |          |          |          |          |
CTTCAAAGAA TATAGATGGT TTTGAAAAGT TCATGCTGTC CCTTCATTGA ATTTTAGAAT    60
GATTGAAGAT AGTGGGAAAA GAGGAAATAC CATGGCAGAA AGAAGACAGC TGTTTGCAGA   120
GATGAGGGCT CAAGATCTGG ATCGCATCCG ACTCTCCACC TACAGAACAG CATGCAAGCT   180
TAGGTTTGTT CAGAAGAAAT GCAATTTGCA CCTGGTGGAC ATATGGAATG TCATAGAAGC   240
ATTGCGGGAA AATGCTCTGA ACAACCTGGA CCCAAACACT GAACTCAACG TGTCCCGCTT   300
AGAGGCTGTG CTCTCCACTA TTTTTTACCA GCTCAACAAG CGGATCACCA CCACTCACCA   360
AATCCATGTG GAGCAGTCCA TCAGCCTCCT CCTTAACTTC CTGCTTGCAG CGTTTGATCC   420
GGAAGGCCAT GGTAAAATTT CAGTATTTGC TGTCAAAATG GCTTTAGCCA CATTGTGTGG   480
AGGGAAGTC ATGGACAAAT TAAGATATAT TTTCTCAATG ATTTCTGACT CCAGTGGGGT   540
GATGGTTTAT GGACGATATG ACCAATTCCT TCGGGAAGTT CTCAAACTAC CCACGGAAGT   600
TTTGGAGGGT CCTTCATTTG GTTACACAGA ACAGTCAGCC AGATCCTGTT TCTCCCAACA   660
GAAAAAAGTC ACGTTAAATG GTTTCTTGGA CACGCTTATG TCAGATCCTG CCCCGCAGTG   720
TCTGGTCTGG TTGCCTCTTC TGCATCGACT AGCAAATGTG GAAAATGTCT TCCATCCGGT   780
TGAGTGTTCC TACTGCCACA GTGAGAGTAT GATGGGATTT CGCTACCGAT GCCAACAGTG   840
TCACAATTAC CAGCTCTGTC AGGACTGCTT CTGGAGGGGA CATGCCGGTG GTTCTCATAG   900
CAACCAGCAC CAAATGAAAG AGTACACGTC ATGGAAATCA CCTGCTAAGA AGCTGACTAA   960
TGCATTAAGC AAGTCCCTGA GCTGTGCTTC CAGCCGTGAA CCTTTGCACC CCATGTTCCC  1020
AGATCAGCCT GAGAAGCCAC TCAACTTGGC TCACATGGTT GATACTTGGC CTCCCAGACC  1080
TGTAACCAGC ATGAACGACA CCCTGTTCTC CCACTCTGTT CCCTCCTCAG GAAGTCCTTT  1140
TATTACCAGG AGCATGCTTG AGAGTTCAAA CCGGCTTGAT GAAGAACACA GGCTAATTGC  1200
CAGGTATGCG GCAAGGCTGG CAGCAGAGTC CTCTTCGTCT CAGCCACCTC AGCAGAGAAG  1260
TGCTCCTGGA ATCTCTTTCA CCATCGATGC GAATAAGCAG CAAAGGCAGC TGATTGCTGA  1320
GCTAGAAAAC AAGAACAGGA AAATCTTACA GGAGATCCAG AGACTTCGGC TAGAGCATGA  1380
ACAAGCTTCT CAGCCCACGC CAGAGAAGGC ACAGCAAAAC CCCACCCTGC TGGCAGAACT  1440
CCGGCTCCTC AGACAGCGCA AAGATGAGCT GGAACAGAGA ATGTCTGCTC TCCAGGAGAG  1500
CCGGAGAGAG CTAATGGTCC AGTGGGAGGG TCTCATGAAG CTACTAAAGG AAGAAGAACT  1560
GAAGCAGGGA GTAAGTTATG TCCCCTACTG CAGGTCTTAA CTAACAGTGG AGGGGCCTGC  1620
CGACCTGCGG TTTTCTCATT GCTTTTGCTC TAATGTATGT TCATGCTTCA GTTTGGAAAG  1680
AGAAAAAAGT CATACTAATT TGCTTCTTTT TCAATGTAGT GCTTGAATTG AGATATATAA  1740
ATTTAGCATT TTTTATAACT ATCACTACTA TCCACATCAA AAGAAGAACT ATGACATCTT  1800
TTAGAAAAGG GAACGAATTG TCATTTATTG GAAACATTTT AGATCCCCAG AGGTATAAGT  1860
TTCAAACCAG TCTTAGCTTT TCAAGTTGTT GATCAGACCC TTCTCTTAAC AGAGAGATAC  1920
CACAGTCACT AGAGATACCC TGAGGTTCAT GTCATCCCAA AACCCACAGC ACTCAGAAGC  1980
TAACCTCTAC ACCCACTCAC ACTGTGAGTA TTCAGTTCGG TTTCATTTTA CTGAAAACCT  2040
GTGAAACCTC TTTTTATAAA AATCAGGCAA TTAAATCCCT TTTCATCACA CAATTATTGA  2100
GCCTTGTTCC CCATGGCTCA CCAAAATGTG CTCAATTTTG TGAGAGAAAG ACTGTACTCC  2160
ATAACTGACT ATTCACGTCC CATCTTTTTG GCTCTTCCCC AAAGCAGAAT CCTTACTGTT  2220
GGTTGACAGT AATCTCTTTT TAAAAAGTAA CTCTCAGCTT TTTCCTTAGC ACCAGAGCCT  2280
TTCGGCTCCG GGAGACGAGA GGGTCATTAC ATACTTTTTT TTTTTTCTGG AAATAGGGGC  2340
ATTGTGACTT TATAGCCTAA ACTGGAGCTG TCTGAACCTG TGGTCAGGCT CAAGAGCCAG  2400
CAGGGGGAGC AGCAAACTC
```

Seq ID NO: 268 Protein sequence
Protein Accession #: AAC50426.1

```
1          11         21         31         41         51
|          |          |          |          |          |
```

```
MIEDSGKRGN  TMAERRQLFA  EMRAQDLDRI  RLSTYRTACK  LRFVQKKCNL  HLVDIWNVIE   60
ALRENALNNL  DPNTELNVSR  LEAVLSTIFY  QLNKRMPTTH  QIHVEQSISL  LLNFLLAAFD  120
PEGHGKISVP  AVKMALATLC  GGKIMDKLRY  IFSMISDSSG  VMVYGRYDQF  LREVLKLPTE  180
VLEGPSFGYT  EQSARSCFSQ  QKKVTLNGFL  DTLMSDPPPQ  CLVWLPLLHR  LANVENVFHP  240
VECSYCHSES  MMGFRYRCOQ  CHNYQLCQDC  FWRGHAGGSH  SNQHQMKEYT  SWKSPAKKLT  300
NALSKSLSCA  SSREPLHPMF  PDQPEKPLNL  AHIVDTWPPR  PVTSMNDTLF  SHSVPSSGSP  360
PITRSMLESS  NRLDEEHRLI  ARYAARLAAE  SSSSQPPQQR  SAPDISFTID  ANKQQRQLIA  420
ELENKNREIL  QEIQRLRLEH  EQASQPTPEK  AQQNPTLLAE  LRLLRQRKDE  LEQRMSALQE  480
SRRELMVQLE  GLMKLLKEEE  LKQGVSYVFY  CRS
```

Seq ID NO: 269 DNA sequence
Nucleic Acid Accession #: NM_001276.1
Coding sequence: 127..1278

```
1          11         21         31         41         51
|          |          |          |          |          |
AGTGGAGTGG  GACAGGTATA  TAAAGGAAGT  ACAGGGCCTG  GGGAAGAGGC  CCTGTCTAGG    60
TAGCTGGCAC  CAGGAGCCGT  GGGCAAGGGA  AGAGGCCACA  CCCTGCCCTG  CTCTGCTGCA   120
GCCAGAATGG  GTGTGAAGGC  GTCTCAAACA  GGCTTGTGG   TCCTGGTGCT  GCTCCAGTGC   180
TGCTCTGCAT  ACAAACTGGT  CTGCTACTAC  ACCAGCTGGT  CCCAGTACCG  GGAAGGCGAT   240
GGGAGCTGCT  TCCCAGATGC  CCTTGACCGC  TTCCTCTGTA  CCCACATCAT  CTACAGCTTT   300
GCCAATATAA  GCAACGATCA  CATCGACACC  TGGGAGTGGA  ATGATGTGAC  GCTCTACGGC   360
ATGCTCAACA  CACTCAAGAA  CAGGAACCCC  AACCTGAAGA  CTCTCTTGTC  TGTCGGAGGA   420
TGGAACTTTG  GGTCTCAAAG  ATTTTCCAAG  ATAGCCTACA  ACACCCAGAG  TCGCCGGACT   480
TTCATCAAGT  CAGTACCGCC  ATTCCTGCGC  ACCCATGGCT  TTGATGGGCT  GGACCTTGCC   540
TGGCTCTACC  CTGGACGGAG  AGACAAACAG  CATTTTACCA  CCCTAATCAA  GGAAATGAAG   600
GCCGAATTTA  TAAAGGAAGC  CCAGCCAGGG  AAAAAGCAGC  TCCTGCTCAG  CGCAGCCACTG  660
TCTGCGGGGA  AGGTCACCAT  TGACAGCAGC  TATGACATTG  CCAAGATATC  CCAACACCTG   720
GATTTCATTA  GCATCATGAC  CTACGATTTT  CATGGAGCCT  GGCGTGGGAC  CACAGGCCAT   780
CACAGTCCCC  TGTTCCGAGG  TCAGGAGGAT  GCAAGTCCTG  ACAGATTCAG  CAACACTGAC   840
TATGCTGTGG  GGTACATGTT  GAGGCTGGCG  GCTCCTGCCA  GTAAGCTGGT  GATGGGCATC   900
CCCACCTTCG  GGAGGAGCTT  CACTCTGGCT  TCTTCTGAGA  CTGGTGTTGG  AGCCCCAATC   960
TCAGGACCGG  GAATTCCAGG  CCGGTTCACC  AAGGAGGCAG  GGACCCTTGC  CTACTATGAG  1020
ATCTGTGACT  TCCTCCGCGG  AGCCACAGTC  CATAGAACCC  TCGGCCCAGCA  GGTCCCCTAT  1080
GCCACCAAGG  GCAACCAGTG  GGTAGGATAC  GACGACCAGG  AAAGCGTCAA  AGCAAGGTG   1140
CAGTACCTGA  AGGATAGGCA  GCTGGCAGGC  GCCATGGTAT  GGGCCCTGGA  CCTGGATGAC  1200
TTCCAGGGCT  CCTTCTGCGG  CCAGGATCTG  CGCTTCCCTC  TCACCAATGC  CATCAAGGAT  1260
GCACTCGCTG  CAACGTAGCC  CTCTGTTCTG  CACACAGCAC  GGGGGCCAAG  GATGCCCCGT  1320
CCCCCTCTGG  CTCCAGCTGG  CCGGGAGCCT  GATCACCTGC  CCTGCTGAGT  CCCAGGCTGA  1380
GCCTCAGTCT  CCCTCCCTTG  GGGCCTATGC  AGAGGTCCAC  AACACACAGA  TTTGAGCCTCA  1440
GCCCTGGTGG  GCAGAGAGGT  AGGGATGGGG  CTGTGGGGAT  AGTGAGGCAT  CGCAATGTAA  1500
GACTCGGGAT  TAGTACACAC  TTGTTGATGA  TTAATGGAAA  TGTTTACAGA  TCCCCAAGCC  1560
TGGCAAGGGA  ATTTCTTCAA  CTCCCTGCCC  CCTAGCCCTC  CTTATCAAAG  GACACCATTT  1620
TGGCAAGCTC  TATCACCAAG  GAGCCAAACA  TCCTACAAGA  CACAGTGACC  ATACTAATTA  1680
TACCCCCTGC  AAAGCCAGCT  TGAAACCTTC  ACTTAGGAAC  GTAATCGTGT  CCCCTATCCT  1740
ACTTCCCCTT  CCTAATTCCA  CAGCTGCTCA  ATAAAGTACA  AGAGTTTAAC  AGTGTGTTGG  1800
CGCTTTGCTT  TGGTCTATCT  TTGAGCGCCC  ACTAGACCCA  CTGGACTCAC  CTCCCCCATC  1860
TCTTCTGGGT  TCCTTCCTCT  GAGCCTTGGG  ACCCCTGAGC  TTGCAGAGAT  GAAGGCCGCC  1920
ATGTT
```

Seq ID NO: 270 Protein sequence
Protein Accession #: NP_001267.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MGVKASQTGF  VVLVLLQCCS  AYKLVCYYTS  WSQYREGDGS  CFPDALDRFL  CTHIIYSFAN   60
ISNDHIDTWE  WNDVTLYGML  NTLKNRNPNL  KTLLSVGGWN  FGSQRFSKIA  YNTQSRRTFI  120
KSVPPFLRTH  GFDGLDLAWL  YPGRRDKQHF  TTLIKEMKAE  FIKEAQPGKK  QLLLSAALSA  180
GKVTIDSSYD  IAKISQHLDF  ISIMTYDFHG  AWRGTTGHHS  PLFRGQEDAS  PDRFSNTDYA  240
VGYMLRLGAP  ASKLVMGIPT  FGRSFTLASS  ETGVGAPISG  PGIPGRFTKE  AGTLAYYEIC  300
DFLRGATVHR  TLGQQVPYAT  KGNQWVGYDD  QESVKSKVQY  LKDRQLAGAM  VWALDLDDFQ  360
GSFCGQDLRF  PLTNAIKDAL  AAT
```

Seq ID NO: 271 DNA sequence
Nucleic Acid Accession #: NM_006474.1
Coding sequence: 181..669

```
1          11         21         31         41         51
|          |          |          |          |          |
GCTGCCTAGG  GTCTGGAAAG  CTCGGGCACC  CTCCCTCTCC  GGGGCTCCTG  CTCCCACCCC    60
TCCGGCCCCC  CCACCGTCGC  GCTCCTCCAG  GCTGGGCCTG  TGGCCGCGGT  GCTTTTAATT   120
TTCCCCCAGC  TCAGAATCTT  GCTGCTCGGC  CCCCAGGAGA  GCAACAACTC  AACGGGAACG   180
ATGTGGAAGG  TGTCAGCTCT  GCTCTTCGTT  TTGGGAAGCG  CGTCGCTCTG  GGTCCTGGCA   240
GAAGGAGCCA  GCACAGGCCA  GCCAGAAGAT  GACACTGAGA  CTACAGGTT   GGAAGGCGGC   300
GTTGCCATGC  CAGGTGCCGA  AGATGATGTG  GTGACTCCAG  GAACCAGCGA  AGACCGCTAT   360
AAGTCTGGCT  TGACAACTCT  GGTGGCAACA  AGTGTCAACA  GTGTAACAGG  CATTCGCATC   420
GAGGATCTGC  CAACTTCAGA  AAGCACAGTC  CACGCGCAAG  AACAAAGTCC  AAGCGCCACA   480
GCCTCAAACG  TGGCCACCAG  TCACTCCACG  GAGAAAGTGG  ATGGAGACAC  ACAGACAACA   540
GTTGAGAAAG  ATGGTTTGTC  AACAGTGACC  CTGGTGGAA   TCATAGTTGG  GGTCTTACTA   600
GCCATCGGTT  TCATTGGTGG  AATCATCGTT  GTGGTTATGC  GAAAAATGTC  GGGAAGGTAC   660
TCGCCCTAAA  GAGCTGAAGG  GTTACGCCCT  GCTTGCCAAC  GTGCTTTAAA  AAAAGACCGT   720
TTCTGACTCT  GTGGCCCTGT  CCCTGAGCTC  GTGGGGAGAA  GATGACCCTG  GGAACATTTG   780
CGGGCCCATT  CAGATTCCAC  GGTGACTTTC  CGTTTGCCAA  ATTAACCGAG  GAAAGACCTT   840
```

TCACCAGATT TGGTTCTTAA ACTTT

Seq ID NO: 272 Protein sequence
Protein Accession #: NP_006465.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MWKVSALLFV LGSASLWVLA EGASTGQPED DTETTGLEGG VAMPGAEDDV VTPGTSEDRY    60
KSGLTTLVAT SVNSVTGIRI EDLPTSESTV HAQEQSPSAT ASNVATSHST EKVDGDTQTT   120
VEKDGLSTVT LVGIIVGVLL AIGFIGGIIV VVMRKMSGRY SP
```

Seq ID NO: 273 DNA sequence
Nucleic Acid Accession #: CAT cluster

```
1          11         21         31         41         51
|          |          |          |          |          |
GCGGCCGCCA GCTTGCAAAG CCGAAGTCTG GCCGCGCTCT TCGACTCGCT GCGCCACGTC    60
CCCGGGGGTG CCGAGCCGGC GGGGGGTGAG GTGGCTGCGC CGGCGGCCGG GCTAGGAGGT   120
GCGGGCACTG GGGGCGCGGG AGGGGCACGTG GCAGGCCCCG CGGGGGCCAC GGCGGATCCCA   180
GGGGCAGGA AGGTCCCGCT GCGGGCACGC AATCTGCCTC CGTCCTTCTT CACGGAGCCG   240
TCCCGGGCAG GCGGGCGGCGG GTGTGGCCCG TCGGGGCAGG ACGTGAGCTT GGGCGACCTG   300
GAGAAGGGCG CGGAGGCCGT GGAGTTCTTT GAGCTGCTGG GGCCCGACTA CGGCGCCGGC   360
ACGGAGGCGG CAGTCTTGCT TGCCGGCCGAG CCTCTCGACG TGTTCCCGCC CGGAGCCTCC   420
GTACTGCGGG GACCCCCGGA GCTGGAGCCC GGCCTCTTTG AGCGCCGCC GGCAGTGGTG   480
GGAAACCTAC TGTACCCCGA GCCCTGGAGC GTCCCGGGCT GCTCCCCGAC CAAAAAGAGC   540
CCCTTGACTG CCCTCCGCGG CGGGGTTGACC TTGAACGAGC CCTTGAGCCC CCTGTACCCC   600
GCCGCTGCGA ATTTCTCCCG GCGGGGAGGA CGGGCCGGGC CATTTGGCTT CTTTCGCCCC   660
CTTCTTTCCA GACTGCGCTT TGC
```

Seq ID NO: 274 DNA sequence
Nucleic Acid Accession #: Eos sequence

```
1          11         21         31         41         51
|          |          |          |          |          |
CAAAGAGGCC GGGCTCCAGC TCCGGGGGTC CCCGCAGTAC GGAGGCTCCG GCGGGGAACA    60
CGTCGAGAGG CTCGGCGGCA AGCAAGACTG CCGCCTCCGT GCCGGCGCCG TAGTCGGGCC   120
CCAGCAGCTC AAAGAACTCC ACGGCCTCCG CGCCCTTCTC CAGGTCGCCC AAGCTCACGT   180
CCGGCCCCGA CGGGCCACAC CCGCCGCCGC CTGCCCGGGA CGGCTCCGTG AAGAAGGACG   240
GAGGCAGATT GCGTGCCCGC AGCGGGACCT TCCTGGCCCC TGGGATCGCC GTGGCCCCCG   300
CGGGGGCCTGC CACGTCCCCT CCCGCGCCCC CAGTGCCCGC ACCTCCTAGC CCGGCCGCCG   360
GCGCAGCCAC CTCACCCCCC GCCGGCTCGG CACCCCCGGG GACGTGGCGC AGCGAGTCGA   420
AGAGCGCGGC CAGACTTCGG CTTTGCAAGC TGGCCGGCCGC
```

Seq ID NO: 275 DNA sequence
Nucleic Acid Accession #: NM_001118.1
Coding sequence: 74..1651

```
1          11         21         31         41         51
|          |          |          |          |          |
AGCCCAGAGA CACATTGGGG CTGACCTGCC GCTGCTGTCA GTGGGAGGCC AGTGGTGCTG    60
GCCAAGAAGT GTCATGGCTG GTGTCGTGCA CGTTTCCCTG GCTGCTCACT GCGGGGCCTG   120
TCCGTGGGGC CGGGGCAGAC TCCGCAAAGG ACGCGCAGCC TGCAAGTCCG CGGCCCAGAG   180
ACACATTGGG GCTGACCTGC CGCTGCTGTC AGTGGGAGCC CAGTGGTGCT GGCCAAGAG   240
TGTCATGGCT GGTGTCGTGC ACGTTTCCCT GGCTGCTCTC CTCCTGCTGC CTATGGCCCC   300
TGCCATGCAT TCTGACTGCA TCTTCAAGAA GGAGCAAGCC ATGTGCCTGG AGAAGATCCA   360
GAGGGCCAAT GAGCTGATGG GCTTCAATGA TTCCTCTCCA GGCTGTCCTG GGATGTGGGA   420
CAACATCACG TGTTGGAAGC CGCCCCATGT GGGTGAGATG GTCCTGGTCA GCTGCCCTGA   480
GCTCTTCCGA ATCTTCAACC CAGACCAAGT CTGGGAGACC GAAACCATTG GAGAGTCTGA   540
TTTTGGTGAC AGTAACTCCT TAGATCTCTC AGACATGGGA GTGGTGAGCC GGAACTGCAC   600
GGAGGATGGC TGGTCGGAAC CCTTCCCTCA TTACTTTGAT GCCTGTGGGT TTGATGAATA   660
TGAATCTGAG ACTGGGGACC AGGATTATTA CTACCTGTCA GTGAAGGCCC TCTACACGGT   720
TGGCTACAGC ACATCCCTCG TCACCCTCAC CACTGCCATG GTCATCCTTT GTCGCTTCCA   780
GAAGCTGCAC TGCACACGCA ACTTCATCCA CATGAACCTG TTTGTGTCGT TCATGCTGAG   840
GGCGATCTCC GTCTTCATCA AAGACTGGAT TCTGTATGCG GAGCAGGACA GCAACCACTG   900
CTTCATCTCC ACTGTGGAAT GTAAGGCGGT CATGGTTTTC TTCCACTACT GTGTTGTGTC   960
CAACTACTTC TGGCTGGTTCA TCGAGGGCCT GTACCTCTTC ACTCTGCTGG TGGAGACCTT  1020
CTTCCCTGAA AGGAGATACT TCTACTGGTA CACCATCATT GGCTGGGGGA CCCCAACTGT  1080
GTGTGTGACA GTGTGGGCTA CGCTGAGACT CTACTTTGAT GACACAGGCGT GCTGGGATAT  1140
GAATGACAGC ACAGCCTCTGT GGTGGGTGAT CAAAGGCCCT GTGGTTGGCT CTATCATGGT  1200
TAACTTGTG CTTTTTATTG GCATTATCGT CATCCTTGTG CAGAAACTTC AGTCTCCAGA  1260
CATGGGAGGC AATGAGTCCA GCATCTACTT GCGACTGGCC CGGTCCACCC TGCTGCTCAT  1320
CCCACTATTC GGAATCCACT ACACACAGTATT TGCCTTCTCC CCAGAGAATG TCAGCAAAAG  1380
GGAAAGACTC GTGTTTGAGC TGGGGCTGGG CTCCTTCCAG GGCTTTGTGG TGGCTGTTCT  1440
CTACTGTTTT CTGAATGGTG AGGTACAAGC GGAGATCAAG CGAAAATGGC GAAGCTGGAA  1500
GGTGAACCGT TACTTCGCTG TGGACTTCAA GCACCGACAC CCGTCTCTGG CCAGCAGTGG  1560
GGTGAATGGG GCCACCCAGC TCTCCATCCT GAGCAAGAGC AGCTCCCAAA TCCGCATGTC  1620
TGGCCTCCCT GCTGACAATC TGGCCACCTG AGCCATGCTC CCCT
```

Seq ID NO: 276 Protein sequence
Protein Accession #: NP_001109.1

```
1          11         21         31         41    -    51
|          |          |          |          |         |
MAGVVHVSLA AHCGACPWGR GRLRKGRAAC KSAAQRHIGA DLPLLSVGGQ WCWPRSVMAG    60
```

```
          VVHVSLAALL LLPMAPAMHS DCIPKKEQAM CLEKIQRANE LMGFNDSSPG CPGMWDNITC   120
          WKPAHVGEMV LVSCPELFRI FNPDQVWETE TIGESDFGDS NSLDLSDMGV VSRNCTEDGW   180
          SEPFPHYFDA CGFDBYESET GDQDYYYLSV KALYTVGYST SLVTLTTAMV ILCRPRKLHC   240
          TRNFIHMNLF VSFMLRAISV FIKDWILYAE QDSNHCFIST VECKAVMVFF HYCVVSNYFW   300
          LFIEGLYLFT LLVETFFPER RYFYWYTIIG WGTPTVCVTV WATLRLYFDD TGCWDMNDST   360
          ALWWVIKGPV VGSIMVNFVL FIGIIVILVQ KLQSPDMGGN ESSIYLRLAR STLLLIPLFG   420
          IHYTVPAFSP ENVSKRERLV FELGLGSFQG FVVAVLYCFL NGEVQAEIKR KWRSWKVNRY   480
          PAVDFKHRHP SLASSGVNGG TQLSILSKSS SQIRMSGLPA DNLAT
```

Seq ID NO: 277 DNA sequence
Nucleic Acid Accession #: NM_004000.1
Coding sequence: 36..1193

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          AGAAGAAGCT GGCCAAGGAT ATGGGAGCAA CCACCATGGA CCAGAAGTCT CTCTGGGCAG   60
          GTGTAGTGGT CTTGCTGCTT CTCCAGGGAG GATCTGCCTA CAAACTGGTT TGCTACTTTA   120
          CCAACTGGTC CCAGGACCGG CAGGAACCAG GAAAATTCAC CCCTGAGAAT ATTGACCCCT   180
          TCCTATGCTC TCATCTCATC TATTCATTCG CCAGCATCGA AAACAACAAG GTTATCATCA   240
          AGGACAAGAG TGAAGTGATG CTCTACCAGA CCATCAACAG TCTCAAAACC AAGAATCCCA   300
          AACTGAAAAT TCTCTTGTCC ATTGGAGGGT ACCTGTTTGG TTCCAAAGGG TTCCACCCTA   360
          TGGTGGATTC TTCTACATCA CGCTTGGAAT TCATTAACTC CATAATCCTG TTTCTGAGGA   420
          ACCATAACTT TGATGGACTG GATGTGAAGCT GGATCTACCC AGATCAGAAA GAAAACACTC   480
          ATTTCACTGT GCTGATTCAT GAGTTAGCAG AAGCCTTTCA GAAGGACTTC ACAAAATCCA   540
          CCAAGGAAAG GCTTCTCTTG ACTGCGGGCG TATCTGCAGG GAGGCAAATG ATTGATAACA   600
          GCTATCAAGT TGAGAAACTG GCAAAAGATC TGGATTTCAT CAACCTCCTG TCCTTTGACT   660
          TCCATGGGTC TTGGGAAAAG CCCCTTATCA CTGGCCACAA CAGCCCTCTG AGCAAGGGGT   720
          GGCAGGACAC AGGGCCAAGC TCCTACTACA ATGTGGAATA TGCTGTGGGG TACTGGATAC   780
          ATAAGGGAAT GCCATCAGAG AAGGTGGTCA TGGGCATCCC CACATATGGG CACTCCTTCA   840
          CACTGCCCTC TGCAGAAACC ACCGTGGGGG CCCCTGCCTC TGGCCCTGGA GCTGCTGGAC   900
          CCATCACAGA GTCTTCAGGC TTCCTGGCCT ATTATGAGAT CTGCCAGTTC CTGAAAGGAG   960
          CCAAGATCAC GCGCCTCCAG GATCAGCAGG TTCCCTACGC AGTCAAGGGG AACCAGTGGG   1020
          TGGGCTATGA TGATGTGAAG AGTATGGAGA CCAAGGTTCA GTTCTTAAAG AATTTAAACC   1080
          TGGGAGGAGC CATGATCTGG TCTATTGACA TGGATGACTT CACTGGCAAA TCCTGCAACC   1140
          AGGGCCCTTA CCCTCTTGTC CAAGCAGTCA AGAGAAGCCT TGGCTCCTTG TGAAGGATTA   1200
          ACTTACAGAG AAGCAGGCAA GATGACCTTG CTGCCTGGGG CCTGCTCTCT CCCAGAGAATT   1260
          CTCATGTGGG ATTCCCCTTG CCAGGCTGGC CTTTGGATCT CTCTTCCAAG CCTTTCCTGA   1320
          CTTCCTCTTA GATCATAGAT TGGACCTGGT TTTGTTTTCC TGCAGCTGTT GACTTGTTGC   1380
          CCTGAAGTAC AATAAAAAAA ATTCATTTTG CTCCAGTA
```

Seq ID NO: 278 Protein sequence
Protein Accession #: NP_003991.1

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          MDQKSLWAGV VVLLLLQGGS AYKLVCYFTN WSQDRQEPGK FTPENIDPFL CSHLIYSFAS   60
          IENNKVIIKD KSEVMLYQTI NSLKTKNPKL KILLSIGGYL FGSKGFHPMV DSSTSRLEFI   120
          NSIILFLRNH NFDGLDVSWI YPDQKENTHF TVLIHELAEA FQKDFTKSTK ERLLLTAGVS   180
          AGRQMIDNSY QVEKLARDLD FINLLSFDFH GSWEKPLITG HNSPLSKGWQ DRGPSSYYNV   240
          EYAVGYWIEK GMPSEKVVMG IPTYGHSFTL ASAETTVGAP ASGPGAAGPI TESSGFLAYY   300
          EICQFLKGAK ITRLQDQQVP YAVKGNQWVG YDDVKSMETK VQFLKNLNLG GAMIWSIDMD   360
          DFTGKSCNQG PYPLVQAVKR SLGSL
```

Seq ID NO: 279 DNA sequence
Nucleic Acid Accession #: NM_015166.1
Coding sequence: 116..1249

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          TGCTGGAAGT CCCTCACCCA GAGACCAGTG CTCCCAACGG CAGAGCAGCG GGGGAGATAA   60
          AGAACTGGTG ACACGTGGCT GTACATTCAG CACAGCTGTG GTGCCCCCAA GTGCCATGAA   120
          CCAGGAGCCA TTCAGAGAGG AGCTGGCCTA TGACCGGATG CCCACGCTGG AGCGGGGCCG   180
          GCAAGACCCC GCCAGCTATG CCCCAGACGC GAAGCCGAGC GACCTGCAGC TGTCGAAGAG   240
          ACTGCCCCCC TGCTTCAGCC ACAAGACGTG GGTCTTCTCT GTGCTGATGG GGAGCTGCCT   300
          CCTGGTGACC TCGGGGTTTT CGCTGTACCT GGGGAACGTG TTCCCGGCTG AGATGGATTA   360
          CTTGCGCTGT GCTGCAGGCT CTTGCATCCC CTCGGCAATT GTGAGCTTCA CGGTCTCCAG   420
          GAGGAACGCC AATGTGATTC CCAACTTTCA GATATTGTTT GTTTCCACGT TTGCTGTGAC   480
          CACTCACGTGT TTAATTTGGT TTGGATGCAA ACTAGTCCTG AACCATCAG CAATAAACAT   540
          CAACTTCAAC CTCATCCTGC TGCTCCTGCT GGAGCTGCTC ATGGCGGCCA CGGTGATCAT   600
          CGGCTGCACGG TCCAGCGAGG AGGACTGCAA GAAAAAGAAG GGCTCCATGT CTGACACGCG   660
          CAACATTCTG GACGAAGTGC CATTTCCTGC TCGGGTCCTG AAATCTTACT CAGTCGTCGA   720
          GGTAATCGCA GGCATCTCTG CCGTCCTCGG GGGGATCATT GCCCTGAACG TGGATGACTC   780
          AGTTTCAGGC CCACACCTCT CAGTGACGTT CTTTTGGATC CTAGTGGCCT GCTTTCCAAG   840
          TGCCATTGCC AGTCATGTGG CAGCAGAGTG TCCCAGCAAG TGTCTGGTGG AGGTCCTGAT   900
          TGCCATAAGC AGCCTCACGT CTCCGCTGCT GTTCACAGCC TCTGGATATC TGTCATTCAG   960
          CATCATGAGA ATCGTGGAGA TGTTTAAGGA TTACCCGCCA GCCATAAAAC CATCCTACGA   1020
          TGTGCTGCTG CTGCTGCTGC TGCTAGTGCT CCTGCTGCAG GCCGGCCTCA ACACGGGCGAC   1080
          CGCCATCCAG TGCGTGCGCT TCAAGGTCAG TGCAAGGCTG CAGGGTGCAT CCTGGGACAC   1140
          CCAGAACGGC CCGCAGGAGC GCCTGGCTGG GGAGGTGGCC AGGAGCCCCC TGAGGGAGTT   1200
          CGACAAGGAG AAAGCCTGGA GAGCCGTCGT GGTGCAAATG CCCCAGACGC CCCAGAATCG   1260
          GGAAACCGGG TGGCAGCGCC CAGCCTGGCC CCAAGCATGG AAACGCACAA CCCCTAAATCG   1320
          CCCTGAGCTA CTGCTTCTAA CACCTCTTTT CCCTTGTGTG AGGGCAAACC AGGCTGCAGG   1380
          TGGGGTTTTC ACTTCCTAGG GTAGTTTAAT TTTAAAATAG GCCAATGTTG GCTAGTCTGT   1440
          GCCTCAGTGA GATCAGTCAG CTCCGAGTGG CTCCCGTGTC GTAACAGCAG GAGCATGGCC   1500
```

1059

```
GCAACTTCCC AGGCCGAGGA AGGGCCCCCG GCTCGGCCTC TTGAGAGCCC CACCCCTGAA   1560
CTGGCCCCAG CTCCTCTTCC TGCCTCTCTC ATGGCTTGGG CTGGAGTGGG CTCTCTGGAC   1620
CTGACCAGAC TGTGGGTCCC TGCGTCTCCT GCCCACTCTG ACCGGGCCTTC CTCCCTCCAC   1680
GCTTAGGGTC TGTCCCGGGT ACTCAGTCAG CCCAGTGGGA TCTTACCCAC TTCCCCTGCAA  1740
GGTGCACCTG CCCCAGGGCTC AGGCTGCCCA GCGGGCTCTTC CTGGACAGTG AGAGCAGGGC  1800
TTGGCGCCTC TGTCCTGGCC CGGGGAGCCGC AGGGGCCCCT CCTCCAGAGC CTGGGCGCAA   1860
GCGACACAGG CTGCCGCTGC TCTCCCAGGT GAAATCCACA CCAGTCCACG CCGGGTCGCC    1920
TGCCCTGTCT CCCTACTTAG ACCCAGTCAT TCTAGAGGGA TCCACCGCCA CACTGGCCGG    1980
CCCACGTCCT GGGTGCTGTC ATGCCCAGCT TGGAGTGCCA CGTGGCCGCT GCCCACGTCC    2040
CGGGCACTGT CATGCCCAGC TTGGAGTGCC ACATGGCCGC TGCCCACGTC CCGGGCACTG   2100
TCATGCCCAG CTTGGAGTGC CACGTGGCCG CTGCTGTGAC AGGCAGTGTT CTTGGGGGTG    2160
GGGCTGCATC CAAGGCTTTG TAAACCGGCT GGACCACGTC TCCCTGGCCC CAGTGACCGG    2220
GGGAAGCTGA GCCCCTCCCT CCTGTGTGTTG CTCCCATTAC TCAAAATGCA GGACAGATCA   2280
GGTCAGAGCC CAGGAATTCT CACAGGTTCA CCCAGCGCCC TCTACCTCCT AGCAAGTACT    2340
TTGTCTTGAT CCTCACTGAG AAGGCCCCAG GGCAGCGGTC TTCTCCATCT CCGCTGTTTT    2400
GGGGTCTTAG GGTACAGCCC AGGCGGTCAC TGCCCACCTG CCAGGCTGCA GGGACAGTTG    2460
GGTGTGAGAA TAACACTGGC TTTGGGTAGT GCCATGGCCA GGAGTGGGTT TCCCTGCGTC    2520
TCCTCGTCCC GAGGGCGCCT GGGTCCTCCC AGCTGACGGC AGTAAATCCA CAGTGAGTTG   2580
GGGCGACTGT GAAACTGGAA TGCTGTTACT TTGATAATTA CTTTCCAGCA GGTGTTTTCC    2640
TTCACAATGG TTTTGTTTCT TTCCTTCTGA TCTGAGAAGA CATGAACGTT TTCTCTTCAC    2700
CGCCGTGGGG TGTATTGACT GGTCCCCCAT GGGCTGCTGG AAAGGCCCGG AGATGCATCT    2760
GTGGCCTGGG GCCATCAAGA TCAAAGAACC AGGAGGCCTG GGAGATGCAG CTGGATGGGG    2820
CGGCCTGCAG ACCCTGCCAG GGGGTTTGAG GACCCTCCCA GGTTTCCCAC TGCGGAACAG    2880
GAGTGACTCT GGCTGCCAAG ATACCTTCAT GGTGTTCATG ACAAGTGGAA TCATTATTTT   2940
CAACCATTGA AGGGGGATGC AGGCAAGACA CCTTCCCAGC TGCTCCTAGA GGGGACAAGC    3000
CAGGCCCTCT CTGCAGTCCT CGGCAGCTCC GGAAGGACAC AGTCAGGGGC CGGGCAAACA    3060
CTTTGGCCAC AGCCCCAAAC AAGCGCCACC GTGGGAGAGG AGAGGCTGCT GTCACTGGTA    3120
CCGGATGCAG ACCCCACCCT GTCTGCAGGC CACCCCCACC TCCCTGCCAGC TTTGAGGCTG    3180
GCGGGGTCTG CTCCTGGGAA TGGGGTGGGA GCCACAGGGA CGACCCGGGG CGGGCTGATG    3240
TCTTCTTGGG GGCAGACCAG AGAGCTCAAG TTTCAGAGTC AGAATTAGGC ACTTGGAACG    3300
TTTTTGCTGG CTTGCACTTT CTTATTTTCT TATTTTAGAG CGCTTAAAAA ATCCGGAAAA    3360
ATGGGGTTTA AAAGAACTGT CTCTTTCAGT CTACATTTTT GTTCAATACG CTTGAGCAAT    3420
AAACGCTGAC TTGCAGACGT G
```

Seq ID NO: 280 Protein sequence
Protein Accession #: NP_055981.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MTQEPFREEL AYDRMPTLER GRQDPASYAP DAKPSDLQLS KRLPPCFSHK TWVFSVLMGS   60
CLLVTSGFSL YLGNVFPAEM DYLRCAAGSC IPSAIVSFTV SRRNANVIPN FQILFVSTFA  120
VTTTCLIWFG CKLVLNPSAI NINFNLILLL LLELLMAATV IIAARSSEED CKKKKGSMSD  180
SANILDEVPF PARVLKSYSV VEVIAGISAV LGGIIALNVD DSVSGPHLSV TFFWILVACF  240
PSAIASHVAA ECPSKCLVEV LIAISSLTSP LLFTASGYLS FSIMRIVEMF KDYPPAIKPS  300
YDVLLLLLL VLLLQAGLNT GTAIQCVRFK VSARLQGASW DTQNGFQERL AGEVARSPLK   360
EFDKEKAWRA VVVQMAQ
```

Seq ID NO: 281 DNA sequence
Nucleic Acid Accession #: NM_004518.1
Coding sequence: 43..2577

```
1          11         21         31         41         51
|          |          |          |          |          |
GCTGAGCCTG AGCCCGACCC GGGGCGCCTC CCGCCAGGCA CCATGGTGCA GAAGTCGCGC    60
AACGGCGGCG TATACCCTGG CCCGAGCGGG GAGAAGAAGC TGAAGGTGGG CTTCGTGGGG   120
CTGGACCCCG GCGCGCCCGA CTCCACCCGG GACGGGGCGC TGCTGATCGC CGGCTCCGAG   180
GCCCCCAAGC GCGGCAGCAT CCTCAGCAAA CCTCGCGCGG GCGGCGCGGG CGCCGGGAAG   240
CCCCCCAAGC GCAACGCCTT CTACCGCAAG CTGCAGAATT TCCTCTACAA CGTGCTGGAG   300
CGGCGCGCCG GCTGGGCGTT CATCTACCAC GCCTACGTGT TCCTCCTGGT TTTCTCCTGC   360
CTCGTGCTGT CTGTGTTTTC CACCATCAAG GAGTATGAGA AGAGCTCGGA GGGGGCCCTC   420
TACATCCTGG AAATCGTGAC TATCGTGGTG TTTGGCGTGG AGTACTTCGT GCGGATCTGG   480
GCCGCAGGCT GCTGCTGCCG GTACCGTGGC TGGAGGGGGC GGCTCAAGTT TGCCCGGAAA   540
CCGTTCTGTG TGATTGACAT CATGGTGCTC ATCGCCTCCA TTGCGGTGCT GGCCGCCGGC   600
TCCCAGGGCA ACGTCTTTGC CACATCTGCG CTCCGGAGCC TGCGCTTCCT GCAGATTCTG   660
CGGATGATCC GCATGGACCG GCGGGGAGGC ACCTGGAAGC TGCTGGGCTC TGTGGTCTAT   720
GCCCACAGCA AGGAGCTGGT CACTGCCTGG TACATCGGCT TCCTTTGTCT CATCCTGGCC   780
TCGTTCCTGG TGTACTTGGC AGAGAAGGGG GAGAACGACC ACTTTGACAC CTACGCGGAT   840
GCACTCTGGT GGGGCCTGAT CACGCTGACC ACCATTGGCT ACGGGGACAA GTACCCCCAG   900
ACCTGGAACG GCAGGCTCCT TGCGGCAACC TTCACCCTCA TCGGTGTCTC CTTCTTCGCG   960
CTGCCTGCAG GCATCTTGGG GTCTGGGTTT GCCCTGAAGG TTCAGGAGCA GCACAGGCAG   1020
AAGCACTTTG AGAAGAGGCG GAACCCGGCA GCAGGCCTGA TCCAGTCGGC CTGGAGATTC   1080
TACGCCACCA ACCTCTCGCG CACAGACCTG CACTCCACGT GGCAGTACTA CGAGCGAAACG   1140
GTCACCGTGC CCATGTACAC ACTTATCCCC CCGCTGAACC AGCTGGAGCT GCTGAGGAAC   1200
CTCAAGAGTA AATCTGGACT CGCTTTCAGG AAGGACCCCC CGCCGGAGCC GTCTCCAAGC   1260
CAGAAGGTCA GTTTGAAAGA TCGTGTCTTC TCCAGCCCCC GAGGCGTGAC TGCCAAGGGC   1320
AAGGGGTCCC CGCAGGCCCA GACTGTGAGG CGGTCACCCA GCGCCGACCA GAGCCTCGAG   1380
GACAGCCCCA GCAAGGTGCC CAAGAGCTGG AGCTTCGGGG ACCGCAGCCG GGCACGCCAG   1440
GCTTTCCGCA TCAAGGGTGC CGCGTCACGG CAGAACTCAG AAGAAGCAAG CCTCCCCGGA   1500
GAGGACATTG TGGATGACAA GAGCTGCCCC TGCGAGTTTG TGACCGGAGGA CCTGACCCCG   1560
GGCCTCAAAG TCAGCATCAG AGCCGTGTGT GTCATGCGGT TCCTGGTGTC CAAGCGGAAG   1620
TTCAAGGAGA GCCTGCGGCC CTACGACGTG ATGGACGTCA TCGACGCAGTA CTCAGCCGGC   1680
CACCTGGACA TGCTGTCCCG AATTAAGAGC CTGCAGTCCA GAGTGGACCA GATCGTGGGG   1740
CGGGGCCCAG CGATCACGGA CAAGGACCGC ACCAAGGGCC GGCCGAGGCC GGAGCTGCCC   1800
GAGGACCCCA GCATGATGGG ACGGCTCGGG AAGGTGGAGA AGCAGGTCTT GTCCATGGAG   1860
```

```
AAGAAGCTGG ACTTCCTGGT GAATATCTAC ATGCAGCGGA TGGGCATCCC CCCGACAGAG   1920
ACCGAGGCCT ACTTTGGGGC CAAAGAGCCG GAGCCGGCGC CGCCGTACCA CAGCCCGGAA   1980
GACAGCCGGG AGCATGTCGA CAGGCACGGC TGCATTGTCA AGATCGTGCG CTCCAGCAGC   2040
TCCACGGGCC AGAAGAACTT CTCGGCGCCC CCGGCCGGCGC CCCCTGTCCA GTGTCCGCCC   2100
TCCACCTCCT GGCAGCCACA GAGCCACCCG CGCCAGGGCC ACGGCACCTC CCCCGTGGGG   2160
GACCACGGCT CCCTGGTGCG CATCCCATGG CCGCCTGCCC ACGAGCGGTC GCTGTCCGCC   2220
TACGGCGGGG GCAACCGCGC CAGCATGGAG TTCCTGCGGC AGGAGGACAC CCCGGGCTGC   2280
AGGCCCCCCG AGGGGACCCT GCGGGACAGC GACACGTCCA TCTCCATCCC GTCCGTGGAC   2340
CACGAGGAGC TGGAGCGGTTC CTTCAGCGGC TTCAGCATCT CCCAGTCCAA GGAGAACCTG   2400
GATGCTCTCA ACAGCTGCTA CGCGGCCGTG GCGCCCTTGTG CCAAAGTCAG GCCCTACATT   2460
GCGGAGGGAG AGTCAGACAC CGACTCCGAC CTCTGTACCC CGTGCGGGCC CCCGCCACGC   2520
TCGGCCACCG GCGAGGGTCC CTTTGGTGAC GTGGGCTGGA CCGGGCCCAG GAAGTGAGGC   2580
GGCGCTGGGC CAGTGGACCC GCCCGCGGCC CTCCTCAGCA CGGTGCCTCC GAGGTTTTGA   2640
GGCGGGAACC CTCTGGGGCC CTTTTCTTAC AGTAACTGAG TGTGGCGGGA AGGGTGGGCC   2700
CTGGAGGGGC CCATGTGGGC TGAAGGATGG GGGCTCCTGG CAGTGACCTT TTACAAAAGT   2760
TATTTTCCAA CAGGGCACTC CCAGGCCCTG TCGCCATTGA GGTGCCTCCG CTGGGCTGTC   2820
TCCTCACCCC TCCTGTGTCT GGAGCCTGTC CCAAAAAGGT GCCAACTGGG AGGCCTCGGA   2880
AGCCACTGTC CAGGCTCCCA CTGCCTGTCT GCTCTGTTCC CAAAGGCAGC GTGTGTGGCC   2940
TCGGGCCCTG CGGTGGCATG AAGCATCCCT TCTGGTGTGG GCATCGCTAC GTGTTTTGGG   3000
GGCAGCGTTT CACGGCGGTG CCCTTGCTGT CTCCCTTGGG CTGGCTCGAG CCTGGGGTCC   3060
ATGTCCCTTT GCCGTCCCGT CATGGGGCAG GGAATCCATA GCGGGGCCCA CAGGCAGGGG   3120
TATGAGTGCG TCCCACCCAA CGCAGCACCA GCCCGGGCCA CCGCTCCCCG TGTCCCCAGT   3180
TCCGTCTCAG CTACCTGGAC TCCAGGACCC TGGAGAAGGG AGACCTGGCA GTGGAGGGAG   3240
GCTGTGCTGT GTGTCCCCCT GCAGGTGTGA CCCCGCCTGC TCTTTCCTCC CCGGCCAGGT   3300
GTGGCCCCGC CTGCTCTTTC CTCCCCCACC AGTATGGCGG CACCTGCTCT TTCCTNCCCC   3360
CCCAAGGTGT GGCCCCACCT GTTCTTTCCT CCCCTGCCGA GGTGTGACCC CACCTGCTCT   3420
TTCCTCCCTC CCAGTATGGC CCCACCTGCT CTTTCCTCCC CGGAGGTGAG GCCCCGCCTG   3480
CTCTTTCCTC CCATGGGAGC CGCTGAGGCG TGCGCACCTG GGCACAGGTT GGGGGCTCTGC   3540
AGGATGAGGA AGACAGGCCA ATCCCTTCCC TCCCAGAAGC TGGCCGCCCA GCAGGAGGGA   3600
CTGAGGCCAG ACTCATGTCC AGCAAGGAAC GTGTGGTGTG TCCCCTGGGA AGTCTCTGGG   3660
CCCTGGGAAG AGGGAAGGTG CACGTCCTGG GATGGTTGCG GGGCCCTGTT TTGGGAGACA   3720
AAGGGGTAGA GGGTCTGTCT TGAGCCCCCC CAGACTCTAG CCCGAGCAGT GCAGCCACCT   3780
ACTGCCCCAC CTCAGAGAAG TGCAGCGGGA AGGAGGCTGG AGGTGGTGCG GCGCTGCCTC   3840
GGGTGTCTGC GTGAATGAGC GTGGCCAAGG ACCAGTGCCA CCTCATGGCA AAGAGCTCCC   3900
GCAGTGTTTG TTAGAGTGCA CATCCTACGT GCCCACTGGC ACACACACGT GCTCACATAC   3960
ANGTCNGNGT ACAGGNGTAC ACATGCANGC TTGCACACAT GCACACAGAC CACATAGCAC   4020
ACATGTGCAC TGACCACACC TGTATAGACC ATGCACAGTA CACATACGTG CATACACATG   4080
CCTGCATACA GGCATACACA TGCACGCTTA CATGTACACG TGCACAGATC ACACACATGC   4140
ACACACGTGT AGCTCACACA CAGTATACAC ATACACAAGT GCACAGACCA CACACAGCAC   4200
TAACACATGC ACACACAAAG TGCATAGGCC ACACAGCACA TGCACACAGG TGCACAGACC   4260
ACACAGCACA CACAAGTGCA CAGAGCACAC TGCACACATG CACACACACA CGCGTGCATG   4320
CACACTCCTC GCACTTCCAG CCTTGGAGCC CTTCTGTCTC TGGTCTTTCT CTTTGACCCT   4380
GCTGAGTGTA AGCTGCCTGG GGAGGGGCTA CAAGGAGTAA TTGTGGCTTT AGGGGTCGTG   4440
GTGATGCTGG AATGTCAAGC GCCGTCGTGG GGTATCCGAC TGTCCGGGCT CCTGGTCCGC   4500
AGTGGCAGAG CGCCAGGCAG AGCCAATCAG GGTCTCGTCG TGCCCTTCCC CCCCACAGCC   4560
TGGCAGCCAT CCAGAGGAGG GGCTCTACCA GATGCCAAGG TGCCCCGGTG TCTGTATGGG   4620
TGTCCGGTTG GGTCCGTGTGT TTGGTCTGCC CTGGAGGTGG CTGGGCCCTC CTGGGATGGG   4680
TGGCTCAGCC TCGAATCCCA GGCCCCAGCC CAGGCAGGTG CTGCTGCCTG TTGTGGTTTC   4740
CTGGCCCAGC TTCTCCTTCT CCCTCTGCAT AAAATCACAG TCCGTGAGTC TTCCAGCTGC   4800
CACCACGGCT GGGACACGCT GGGGGAGGGC TCCTCCCATG CCTCCTGCAC ACAGCCGTCT   4860
·GAGCAGGGCA GGTGCCAACA CCCCCCACCG GAGACACGCT GCCCCTCAGC GATGCCCCTA   4920
CCTTTTGGGG GGCCTCGTCT CAAGCCCCCC CTTGGAGGCT GAAATCACCC CAGGCACTGT   4980
GAGGGCTTCT CCAGGGGGAC ACCCTTTGAG CTGTGGGTCT GATCACCCCA AGTCCCGCAC   5040
ACGGAGGAGA GGCACAGCCA GGGCGTGTGG TTTAATGTTT GCCCCTTCGG GGCTGGAGGT   5100
CTCAGTGTTT CTAGATTCCA GACCCTGCTG CCAGAGAGAC CTGCTGCCGG AGAGAAGGGG   5160
AGGAGGACTC CAGCTGGGCT CGGTCCCCCA CAGTCAGGGA CCCCCATAAA GGACACCCCC   5220
TTCTCTCTAG AAAGAGCTGG GCTCTCAGCT ATTTCTAGTT GCTTCCCAGA AGCCGAGGAG   5280
CAGAAGGAGC TGTGAGAGCT TTGCAGAAAC GCCCTTGTCC CGCCCTCCT GAGCTATGAA   5340
TGCCGTACAG AGCAGAGGCT GGGGCATTGG CAAGATCACA GGTTGATGCT GCACAGCCCC   5400
ATTGACACAA ACCCTCAAAG CAGACGTGAG AGGGACGGTT CACAAAGCTT GGACCTGCCG   5460
TGGAGGGTGC CCGGCAGACG TGGCGTGAGA GGGACGGCTC ACGAGGCTTG GACCTGCTGT   5520
GGAGGGTGCC CAGCAGACGT GGTGTGAGAG GAACGGCTCA CGAGACTTGG ACCTGGTGGA   5580
GGGTGCCCAG CAGACGTGGT GTGAGAGGGA CGGCTCACAG GGCTTGGACC GGAGAGAGAT   5640
GGCTCATGAG ACTTGGACCT GCCGTGGAGG GTGCCCAGCA GACGTGGTAT GAGAGGGATG   5700
GCTCACGAGG CTTGGACCTG GTGGAGGGTG CCCGGCAGAC GTGTGAGAGG GACGGTTCAC   5760
AAGGCTTGGA CCTGCCATGG AGGGTGCCCA GCAGACGTGG TGTGAGAGGG ACAGCTCACG   5820
AGGCTTGGAC CTGCCGTGGA GGGTGCCCAG CAGGGGGCTG AGCTCTGAGG GGTGGGTGCT   5880
CAGTGCACGG GTGCCCCCAG TGTCCTCTGA TCCTGTCCGG TGCCTCCCCC AACCCCCACA   5940
CCCATGCAGA ACTCCCAGGT CACATGCACG TATGTCCAGG GCATGGGGGT GGCGTGAAGA   6000
GGCTGGTCA GGGCCTTTAG GGGCTGCAGG ACGGAATGGC CACCTGGGGA GCCTGTGTGG   6060
CTGTGCCGGG CAGCCATCCT GCATTCCCAC CCAGCGCGCA GTCTCCACCT CGGCCCCAGC   6120
AAAGCGCTAA GCAGCCGGAG AGACAGCCAG GGCGGCTTCC TGAAGGATGT GGGATGGTGG   6180
ACTCCGGGGT CGAGGGAATA CGCAGGTTCC TGTCCTCCGG GAGACCTAGA GAAGCTGCAC   6240
ACCCAGGAGC TTTCCATGAC CCGGGAGCAT GAGTGAATGG GGGGTTCCAG TTTGCTGAAC   6300
TTTGCTGTCT TGTAAGGGTG GGGGCTGACG GCCGACCCTG GGAGGAGGTG ACACCGCAGG   6360
GGGAGGTTGT GGGCAACGGT GGAGGAGGAG AGACGGGAGG GGACCATTTG GGATGGAGGG   6420
GCCTCTTCAG AGTTTTAAAA GGCGTTTGTG GGGTGGAGTT GAGTGTGCTC TGGGCTTGGA   6480
CACTTGCCGT GGTGCCCCTG GCTGGCCGAG GAGACTGGCT CTGGCCAGGG CCCCGTCCTG   6540
AGAGGTCCTC AGCGTCTGAC TCTCGGCCGA GCGCCAGCAA GGAGGGGCCG GTCCCCGGGG   6600
CTACCAGGCA GGCACGTGCA CATCGCCATC GCCACACGCC AACTCGGCCT GGGTTTTACA   6660
AAGTCGTTGC CTTAATGCAT GTGGACAGGA ACTCCCTGAG GTCGCCCCAT GCCCCCTGGC   6720
TGTGCCAGGT ACGGACGCCC TGGACCCTGC GAACAGGTGG GGCGGGCGAG GGGCCCAAGG   6780
GAGCGGGCTCC AGAGACACGC GCAGGGCAGG AGGGGTCTCA CGGAGGGGTC TCGCACTGAG   6840
GCGCCCAGAG CTGGTGGTCC CGCTGGACGC CATCCCTCTG CCCGGGATCC ACACGGCCCA   6900
```

```
CGTGTGCCCG CCATGCCCGC GCCCCACGCC ATTGCAGTCT TCCATCCTCT GGCCGTGACG  6960
GTGGCTGCAG CTTCCCCATT TGCGCCGTTG CCTCTGGCTG TCTGCACTTT TGTTCATGCT  7020
CCAAAGAACA TTTCATAATG CCTTCAGTAC CGACGTACAC TTCTGACCAT TTTGTATGTG  7080
TCCTTGTGCC GTAGTGACCA GGCCTTTTTT TGGTGGATGT GTTACCCCGC ACACTTCAAT  7140
CTCAACTTTG TGCACCGTCC ATTTTCTAGG GATAGACGCC CAGGGAATGA ACTCTAGTTT  7200
TCTAACAGAT TAGCTGAGAT ATTAACTTAC TCACACGGAC AGGTTGATGC CAGAGCCGAT  7260
AGAATGCGCC AGTGCGGGTT TGCGGGGGAC TTCGGGTGTG GGGTCCTGCG GCCGCGATGG  7320
CCGTGGAAGG TTCTGGGGAT CCCTGCTGCC ACGGGGACGA GTTCGGACGC CAGGTGGACC  7380
TGTGCACTCA GTAAAACGCA GTGATTCAAA AAAAAAAAA
```

Seq ID NO: 282 Protein sequence
Protein Accession #: NP_004509.1

```
      1         11        21        31        41        51
      |         |         |         |         |         |
MVQKSRNGGV YPGPSGEKKL KVGFVGLDPG APDSTRDGAL LIAGSEAPKR GSILSKPRAG   60
GAGAGKPPKR NAFYRKLQNF LYNVLERPRG WAFIYHAYVF LLVPSCLVLS VFSTIKEYEK  120
SSEGALYILE IVTIVVFGVE YFVRIWAAGC CCRYRGWRGR LKFARKPFCV IDIMVLIASI  180
AVLAAGSQGN VFATSALRSL RFLQILRMIR MDRRGGTWKL LGSVVYAHSK ELVTAWYIGF  240
LCLILASFLV YLAERGENDH FDTYADALWW GLITLTTIGY GDKYPQTWNG RLLAATFTLI  300
GVSFFALPAG ILGSGFALKV QEQHRQKHFE KRRNPAAGLI QSAWRPYATN LSRTDLHSTW  360
QYYERTVTVP MYRLIPPLNQ LELLRNLKSK SGLAFRKDPP PEPSPSQKVS LKDRVFSSPR  420
GVAAKGKGSP QAQTVRRSPS ADQSLEDSPS KVPKSNSFGD RSRARQAFRI KGAASRQNSE  480
EASLPGEDIV DDKSCPCEFV TEDLTPGLKV SIRAVCVMRF LVSKRKFKES LRPYDVMDVI  540
EQYSAGHLDM LSRIKSLQSR VDQIVGRGPA ITDKDRTKGP AEAELPEDPS MMGRLGKVEK  600
QVLSMEKKLD FLVNIYMQRM GIPPTETEAY FGAKEPEPAP PYHSPEDSRE HVDREGCIVK  660
IVRSSSSTGQ KNFSAPPAAP PVQCPPSTSW QPQSHPRQGH GTSPVGDEGS LVRIPPPPAH  720
ERSLSAYGGG NRASMEFLRQ EDTPGCRPPE GTLRDSDTSI SIPSVDHEEL ERSFSGFSIS  780
QSKENLDALN SCYAAVAPCA KVRPYIAEGE SDTDSDLCTP CGPPPRSATG EGPFGDVGWA  840
GPRK
```

Seq ID NO: 2838 DNA sequence
Nucleic Acid Accession #: AF152496.1
Coding sequence: 1..2391

```
      1         11        21        31        41        51
      |         |         |         |         |         |
ATGGAGGCGG GAGGAGAGCG ATTTCTTAGA CAAAGGCAAG TCTTGCTTCT CTTTGTTTTT   60
CTGGGGAGGGT CTCTGGCTGG GTCCGAGTCA AGAGCGCTATT CTGTGGCTGA GGAAAAAGAG  120
AAGGGCTTTT TAATAGCCAA CCTAGCAAAG GATCTGGGAC TAAGGGTAGA GGAACTGGCC  180
GCGAGGGGGG CCCAAGTTGT GTCCAAAGGG AACAAACAGC ATTTTCAGCT CAGTCATCAG  240
ACAGGTGATT TGCTCCTGAA TGAGAAATTG GACCGGGAGG AGCTATGCGG CCCCACAGAA  300
CCATGCATAC TACATTTTCA GATATTACTG CAAAACCCTT TGCAATTCGT TACAAACGAG  360
CTCCGTATCA TAGATGTAAA TGACCATTCT CCGGTATTCT TTGAAAATGA AATGCATCTG  420
AAAATCCTAG AAAGCACTCT GCCAGGAACA GTAATTCCTT TGGGAAATGC TGAGGACTTG  480
GATGTGGGAA GAAACAGCCT CCAAAACTAC ACTATCACTC CGAATTCCCA CTTCCACGTA  540
CTCACTCGCA GTCGTAGGGA CGGAAGGAAG TACCCGGAAC TAGTACTGGA TAAAGCGTCA  600
GATCGGGAGG AGCAGCCGGA ACTCAGCTTA ACGCTCACCG CGCTGGACGG CGGCTCTCCC  660
CCTCGGTCTG GGACAGCCCA GATAAACATC CAGGTCTTAG ATATAAACGA CAATGCACCA  720
GAATTTGCAC AGCCGCTCTA TGAGGTTGCA GTTCTAGAGA ATACCCCCGT TAACTCTGTC  780
ATTGTCACTG TCTCGGCTTC TGACTTAGAT ACAGGAAGTT TTGGGACAAT ATCATATGCA  840
TTTTTTCATG CTTCTGAAGA AATTGCAAA ACTTTTCAGC TAAATCCAAT TACTGGTGAT  900
ATGCAACTGG TCAAATATTT GAATTTTGAA GCGATTATCA GTTATGAAGT CGACATCGAG  960
GCCAAGGATG GCGGAGGCCT ATCCGGAAAG TCTACAGTCA TAGTCCAGGT GGTTGATGTC 1020
AACGACAACC CACCGGAACT GACCTTGTCT TCAGTAAACA GCCCTATTCC TGAGAACTCG 1080
GGAGAGACTG TACTGGCTGT TTTCAGTGTT TCTGATCTAG ACTCTGGAGA CAACGGAAGA 1140
GTGATGTGTT CCATTGAGAA CAATCTCCCC TTCTTCCTGA AACCATCTGT AGAGAATTTT 1200
TACACCCTAG TGTCAGAAGG CGCGCTGGAC AGAGAGACCA GATCCGAGTA CAACATTACC 1260
ATCACTATCA CTGACCTGGG GACACCCAGG CTGAAAACCA AGTACAACAT AACCGTGCTG 1320
GTCTCCGACG TCAATGACAA CGCCCCCGCC TTCACCCAAA TCTCCTACAC CCTGTTCGTC 1380
CGCGAGAACA ACAGCCCCGC CCTGCACATC GGCAGTGTCA GCGCCACAGA CAGAGACTCA 1440
GGCACCAACG CCCAGGTAAC CTACTCGCTG CTGCCGGCCC AGGACCCGCA CCTGCCCCTC 1500
TCTTCCCTGG TCTCCATCAA CGCGGACAAC GGCCACCTGT TTGCCCTCAG GTCGCTGGAC 1560
TACGAGGCCC TGCAGGCGTT CGAGTTCCGC GTGGGCGCCA CAGACCGTGG CTCCCCGGCT 1620
TTGAGCAGCG AGGCGCTGGT GCGCGTGCTG GTGCTGGACG CCAACGACAA CTCGCCCTTC 1680
GTGCTGTACC CGCTGCAGAA CGGCTCCGCG CCCTGCACCG AGCTGGTGCC CCGGGCGGCT 1740
GAGCCGGGCT ACTGGTGAC CAAGGTGGTG GCGGTGGACG GCGACTCGGG CCAGAACGCC 1800
TGGCTGTCGT ACCAGCTGCT CAAGGCCACG GAGCCCGGC TGTTCGGCGT GTGGGCGCAC 1860
AATGGCGAAG TGCGCACCGC CAGGCTGCTG AGCGAGCGGG ACGCGGCCAA GCACAGGCTG 1920
GTGGTGCTGG TCAAGGACAA TGGCCAGCCT CCGCGCTCGG CCACCGCCAC GCTGCATGTG 1980
CTCCTGGTGG ACGGCTTCTC CCAGCCCTAC CTGCCTCTCC CGGAGGCGGC ACCGGCCCAG 2040
GCCCAGGCCG ACTTGCTCAC CGTCTACCTG GTGGTGGCAT TGGCCTCGGT GTCTTGCCTC 2100
TTCCTCTTTT CGGTGCTCCT GTTCGTGGCG GTGCGGCTGT GCAGGAGGAG CAGGGCGGCC 2160
TCGGTGGGTC GCTGCTCGGT GCCCGAGGGC CCCTTTCCAG GGCAGATGGT GGACGTGAGC 2220
GGCACCGGGA CCCTGTCCCA GAGCTACCAG TACGAGGTGT GTCTGACTGG AGGCTCCGGG 2280
ACAAATGAGT TCAAGTTCCT GAAGCCAATT ATCCCCAACT TCGTTGCTCA GGGTGCAGAG 2340
AGGGTTAGCG AGGCAAATCC CAGTTTCAGG AAGAGCTTTG AATTCAGTTA A            
```

Seq ID NO: 284 Protein sequence
Protein Accession #: AAD43757.1

```
      1         11        21        31        41        51
      |         |         |         |         |         |
MEAGGERFLR QRQVLLLFVF LGGSLAGSES RRYSVABEKE KGFLIANLAK DLGLRVBELA   60
```

```
        ARGAQVVSKG NKQHFQLSEQ TGDLLLNEKL DREELCGPTE PCILHFQILL QNPLQFVTNE    120
5       LRIIDVNDHS PVFPENEMHL KILESTLPGT VIPLGNAEDL DVGRNSLQNY TITPNSHPHV    180
        LTRSRRDGRK YPELVLDKAL DREEQPELSL TLTALDGGSP PRSGTAQINI QVLDINDNAP    240
        EFAQPLYEVA VLENTFVNSV IVTVSASDLD TGSFGTISYA FFHASEEIRK TFQLNPITGD    300
        MQLVKYLNPE AINSYEVDIE AKDGGGLSGK STVIVQVVDV NDNPPELTLS SVNSPIPENS    360
25      GETVLAVFSV SDLDSGDNGR VMCSIENNLP PPLKPSVENF YTLVSEGALD RETRSEYNIT    420
        ITITDLGTPR LKTKYNITVL VSDVNDNAPA FTQISYTLFV RENNSPALHI GSVSATDRDS    480
        GTNAQVTYSL LPPQDPHLPL SSLVSINADN GHLPALRSLD YEALQAFEFR VGATDRGSPA    540
        LSSEALVRVL VLDANDNSPP VLYPLQNGSA PCTELVPRAA EPGYLVTKVV AVDGDSGQNA    600
10      WLSYQLLKAT EPGLFGVWAH NGEVRTARLL SERDAAKHRL VVLVKDNGEP PRSATATLHV    660
        LLVDGPSQPY LPLPEAAPAQ AQADLLTVYL VVALASVSSL FLFSVLLFVA VRLCRRSRAA    720
        SVGRCSVPEG PFPGQMVDVS GTGTLSQSYQ YEVCLTGGSG TNEFKFLKPI IPNFVAQGAE    780
        RVSEANPSFR KSFEFS
```

15   Seq ID NO: 285 DNA sequence
Nucleic Acid Accession #: NM_001794.2
Coding sequence: 15..2765

```
20      1          11         21         31         41         51
        |          |          |          |          |          |
        CGCCGGCGGG GAAGATGACC GCGGGCGCCG GCGTGCTCCT TCTGCTGCTC TCGCTCTCCG    60
        GCGCGCTCCG GGCCCATAAT GAGGATCTTA CAACTAGAGA GACCTGCAAG GCTGGGTTCT    120
        CTGAAGATGA TTACACGGCA TTAATCTCCC AAAATATTCT AGAAGGGGAA AAGCTACTTC    180
        AAGTCAAGTT CAGCAGCTGT GTGGGGACCA AGGGGACACA ATATGAGACC AACAGCATGG    240
25      ACTTCAAAGT TGGGGCAGAT GGGACAGTCT TCGCCACCCG GGAGCTGCAG GTCCCCTCCG    300
        AGCAGGTGGC GTTCACGGTG ACTGCATGGG ACAGCCAGAC AGCAGAGAAA TGGGACGCCG    360
        TGGTGCGGTT GCTGGTGGCC CAGAACCTGT CCCCGCACTC TGGACACAAG CCGCAGAAAG    420
        GAAAGAAGGT CGTGGCTCTG GACCCCTCTC CGCCTCCGAA GGACACCCTG CTGCCGTGGC    480
        CCCAGCACCA GAACGGCCAAC GGGCTGAGGC GGCGCAAACG GGACTGGGTC ATCCCGCCCA    540
30      TCAACGTGCC CGAGAACTCG CGCGGGCCCT TCCCGCAGCA GCTCGTGAGG ATCCGGTCCG    600
        ACAAAGACAA TGACATCCCC ATCCGGTACA GCATCACGGG AGTGGGCGCC GACCAGCCCC    660
        CCATGGAGGT CTTCAGCATT GACTCCATGT CCGGCCGGAT GTACGTCACA AGGCCCATGG    720
        ACCGGGAGGA GCACGCCTCT TACCACCTCC GAGCCCACGC TGTGGACATG AATGGCAACA    780
        AGGTGGAGAA CCCCATCGAC CTGTACATCT ACGTCATCGA CATGAATGAC AACCGCCCTG    840
35      AGTTCATCAA CCAGGTCTAC AACGGCTCCG TGGACGAGGG CTCCAAGCCA GGCACCTACG    900
        TGATGACCGT CACGGCCAAC GATGCTGACG ACAGCACCAC GGCCAACGGG ATGGTGCGGT    960
        ACCGGATCGT GACCCAGACC CCACAGAGCC CGTCCCAGAA TATGTTCACC ATCAACGCCG    1020
        AGACTGGAGA TATCGTCACA GTGGCGGCTG GCCTGGACCG AGAGAAAGTT CAGCAGTACA    1080
        CAGTCATCGT TCAGGCCACA GATATGGAAG GAAATCTCAA CTATGGCCTC TCAAACACAG    1140
40      CCACAGCCAT CATCACGGTG ACAGATGTGA ATGACAACCC GCCAGAATTT ACCGGCAGCA    1200
        CGTTTGCAGG GGAGGTCCCC GAAAACCGCG TGGAGACCGT GGTCGCAAAC CTCACGGTGA    1260
        TGGACCGAGA TCAGCCCCAC TCTCCAAACT GGAATGCCGT TTACCGCATC ATCAGTGGGG    1320
        ATCCATCCGG GCACTTCAGC GTCCGCACAG ACCCCGTAAC CAACGAGGGC ATGGTCACCG    1380
        TGGTGAAGGC AGTCGACTAC GAGCTCAACA GAGCTTTCAT GCTCGACAGTG ATGGTGTCCA    1440
45      ACCAGGCCGC CCTGGCCAGC GGAATCCAGA TGTCCTTCCA GTCCACGGCA GGGGTGACCA    1500
        TCTCCATCAT GGACATCAAC GAGGCTCCCT ACTTCCCCTC AAACCACAAG CTGATCCGCC    1560
        TGGAGGAGGG CGTGCCCCCC GGCACCGTGC TGACCACGTT TTCAGCTGTG GACCCTGACC    1620
        GGTTCATGCA GCAGGCTGTG AGATACTCAA AGCTGTCAGA CCCAGCGAGC TGGCTGCACA    1680
        TCAATGCCAC CAACGGCCAG ATCACCACGG CGGCAGTGCT GGACCGTGAG TCCCTCTACA    1740
50      CCAAAAACAA CGTCTACGAG GCCACCTTCC TGGCAGCTGA CAATGGGATA CCCCCGGCCA    1800
        GCGGCACCGG GACCCTCCAG ATCTATCTCA TTGACATCAA CGACAACGCC CCTGAGCTGC    1860
        TGCCCAAGGA GGCGCAGATC TGCGAGAAGC CCAACCTGAA CGCCATCAAC ATCACGGCGG    1920
        CCGACGCTGA CGTCGACCCC AACATCGGCC CCTACGTCTT CGAGCTGCCC TTTGTCACGG    1980
        CGGCCGTGCG GAAGAACTGG ACCATCACCG GCCTGAACGG TGACTATGCC CAACTCAGCT    2040
55      TGCGCATCCT GTACCTGGAG GCCGGGATGT ATGACGTCCC CATCATCGTC ACAGACTCTG    2100
        GAAACCCTGC CCTGTCCAAC AGTCCATCA TCAAAGTCAA GGTGTGCCCA TGTGATGACA    2160
        ACGGGGACTG CACCACCATT GGCGCAGTGG CAGCGGCTGG TCTGGGCACC GGTGCCATCG    2220
        TGGCCATCCT CATCTGCATC CTCATCCTGC TGACCATGGT CCTGCTGTTT GTCATGTGGA    2280
        TGAAGCGGCG AGAGAAGGAG CGCCACACGA AGCAGCTGCT CATTGACCCC GAGGACGACG    2340
60      TCGGCGACAA CATCCTCAAG TATGACGAGG AAGGCGGTGG CGAGGAGGAC CAGGACTACG    2400
        ACCTCAGCCA GCTGCAGCAG CCGGAAGCCA TGGGGCACGT GCCAAGCAAA GCCCCTGGCG    2460
        TGCGTCGGGT GGATGAGCGG CCGGTGGGCG CTGAGCCCCA GTACCGGATC AGGCCCATGG    2520
        TGCCGCACCC AGGCGACATC GGTGACTTCA TCAATGAGGG ACTCCGCGCT GCTGACAACG    2580
        ACCCCACGGC ACCCCCCTAT GACTCCCTGC TGGTCTTCGA CTACGAGGGG AGCGGCTCCA    2640
65      CCGCAGGCTC CGTCAGCTCC CTGAACTCAT CCAGATTCCGG GGACCAAGAC TACGATTACC    2700
        TCAACGACTG GGGGCCCAGA TTCAAGAAGC TGGCGGACAT GTATGGAGGT GGTGAAGAGG    2760
        ATTGACTGAC CTCGCATCTT CGGACTGAAG TGAGAGCCGT GCTCGGACGC CGGAGGAGCA    2820
        GGACTGAGCA GAGGCGGCCG GTCTTCCCGA CTCCCTGCGG CTGTGTCCTT AGTGCTGTTA    2880
        GGAGGCCCCC CAATCCCCAC GTTGAGCTGT CTAGCATGAG CACCCACCCC CACACGCGCCC    2940
70      TGCACCCGGC CGCTGCCCAG CACCGCGCTG GCTGGCACTG AAGGACAGCA AGAGGCACTC    3000
        TGTCTTCACT TGAATTTCCT AGAACAGAAG CACTGTTTTT AAAAAAAAAA AAAAAAAAG    3060
        AAG
```

Seq ID NO: 286 Protein sequence
Protein Accession #: NP_001785.2

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        MTAGAGVLLL LLSLSGALRA HNEDLTTRET CKAGFSEDDY TALISQNILE GEKLLQVKFS    60
        SCVGTKGTQY ETNSMDFKVG ADGTVFATRE LQVPSEQVAF TVTAWDSQTA EKWDAVVRLL    120
        VAQTSSPHSG HKPQKGKKVV ALDPSPPPKD TLLPWPQHQN ANGLRRRKRD WVIPPINVPE    180
        NSRGPFPQQL VRIRSDKDND IPIRYSITGV GADQPPMEVF SIDSMSGRMY VTRPMDREEH    240
        ASYHLRAHAV DMNGNKVENP IDLYIYVIDM NDNRPEFINQ VYNGSVDEGS KPGTYVMTVT    300
        ANDADDSTTA NGMVRYRIVT QTPQSPSQNM FTINSETGDI VTVAAGLDRE KVQQYTVIVQ    360
```

```
      ATDMEGNLNY GLSNTATAII TVTDVNDNPP EFTASTFAGE VPENRVETVV ANLTVMDRKQ    420
      PHSPNWNAVY RIISGDPSGH FSVRTDPVTN EGMVTVVKAV DYELNRAFML TVMVSNQAPL    480
      ASGIQMSFQS TAGVTISIMD INEAPYFPEN HKLIRLEEGV PPGTVLTTFS AVDPDRFMQQ    540
      AVRYSKLSDP ASWLHINATN GQITTAAVLD RESLYTKNNV YEATFLAADN GIPPASGTGT    600
      LQIYLIDIND NAPELLPKEA QICEKPNLNA INITAADADV DPNIGPYVFE LPFVPAAVRK    660
      NWTITRLNGD YAQLSLRILY LEAGMYDVPI IVTDSGNPPL SNTSIIKVKV CPCDDNGDCT    720
      TIGAVAAAGL GTGAIVAILI CILILLLTMVL LFVMWMKRRE KERHTKQLLI DPEDDVRDNI    780
      LKYDEEGGGE EDQDYDLSQL QQPEAMGHVP SKAPGVRRVD ERPVGAEPQY PIRPMVPHPG    840
      DIGDFINEGL RAADNDPTAP PYDSLLVFDY EGSGSTAGSV SSLNSSSSGD QDYDYLNDWG    900
      PRFKKLADMY GGGEED
```

Seq ID NO: 287 <u>DNA sequence</u>
Nucleic Acid Accession #: AF152495.1
Coding sequence: 1..2397

```
       1         11        21        31        41        51
       |         |         |         |         |         |
      ATGGAGGCCG GAGAGGGGAA GGAGCGCGTT CCGAAACAAA GGCAAGTCCT GATATTCTTT    60
      GTTTTGCTGG GCATAGCTCA GGCTAGTTGC CAGCCTAGGC ACTATTCAGT GGCCGAGGAA    120
      ACGGAGAGTG GCTCCTTTGT GGCCAATTTG TTAAAGACCC TGGGGCTGGA GATAGGAGAA    180
      CTTGCTGTGA GGGGGGCCAG GGTCGTTTCC AAAGGAAAAA AAATGCATTT GCAGTTCGAT    240
      AGGCAGACCG GGGATTTGTT GTTAAATGAG AAATTGGACC GGGAGGAGCT GTGCGGCCCC    300
      ACAGAGCCCT GTGTCCTACC TTTCCAGGTG TTACTAGAAA ATCCCTTGCA GTTTTTTCAG    360
      GCGGAGCTAC GGATTAGGGA CGTAAATGAT CATTCCCCAG TTTTCCTAGA CAAAGAAATA    420
      CTTTTGAAAA TTCCAGAAAG TATCACTCCT GGAACTACTT TCTTAATAGA ACGTGCCCAG    480
      GACTTGGATG TAGGAACCAA CAGTCTCCAA AATTACACAA TCAGTCCCAA TTTCCACTTT    540
      CATCTTAATT TACAAGACAG TCTCGATGGC ATAAATATTAC CACAGCTGGT GCTGAACAGA    600
      GCCCTGGATC GCGAGGAGCA GCCTGAGATC AGGTTAACCC TCACAGCGCT AGATGGCGGG    660
      AGTCCACCCA GGTCGGCAC GGCCCTGGTA CGGATTGAAG TTGTGGACAT CAATGACAAC    720
      GTCCCAGAGT TTGCAAAGCT GCTCTATGAG GTGCAGATCC CGGAGGACAG CCCCGTTGGA    780
      TCCCAGGTTG CCATCGTCTC TGCCAGGGAT TTAGACATTG GAACTAATGG AGAAATATCT    840
      TATGCATTTT CCCAAGCATC TGAAGACATT CGCAAAACGT TTCGATTAAG TGCAAAATCG    900
      GGAGAACTGC TTTTAAGACA GAAACTGGAT TTCGAATCCA TCCAGACATA CACAGTAAAT    960
      ATTCAGGCGA CAGATGGTGG GGGCCTATCT GGAACTTGTG TGGTATTTGT CCAAGTGATG    1020
      GATTTGAATG ACAATCCTCC GGAACTAACT ATGTCGCACAC TTATCAATCA GATCCCAGAA    1080
      AACTTGCAGG ACACCCTCAT TGCTGTATTC AGCGTTTCAG ATCCTGACTC CGGAGACAAC    1140
      GGAAGGATGG TGTGCTCCAT CCAAGATGAT CTTCCTTTTT TCTTGAAACC TTCTGTTGAG    1200
      AACTTTTACA CTCTGGTGAT AAGCACGGCC CTGGACCGGG AGACCAGATC CGAATACAAC    1260
      ATCACCATCA CCGTCACCGA CTTCGGGACA CCCAGGCTGA AAACCGAGCA CAACATAACC    1320
      GTGCTGGTCT CCGACGTCAA TGACAACGCC CCCGCCTTCA CCCAAACCTC CTACACCCTG    1380
      TTCGTCCGCG AGAACAACAG CCCCGCCCTG CACATCGGCA GCGTCAGCGC CACAGACAGA    1440
      GACTCGGGCA CCAACGCCCA GGTCACCTAC TCGCTGCTGC CGCCCCAGGA CCCGCACCTG    1500
      CCCCTCGCCT CCCTGGTCTC CATCAACGCG GACAACGGCC ACCTGTTCGC TCTCCAGTCG    1560
      CTGGACTACG AGGCCCTGCA GGCGTTCGAG TTCCGCGTGG GCGCCGCAGA CCGCGGCTCC    1620
      CCGGCGTTGA GCAGCGAGGC GCTGGTGCGC GTGCTGGTGC TGGACGCCAA CGACAACTCG    1680
      CCCTTCGTGC TGTACCCGCT GCAGAACGGC TCCGCGCCCT GCACCGAGCT GGTGCCCCGG    1740
      GCGGCCGAGC CGGGCTACCT GGTGACCAAG GTGGTGGCGG TCGGGAGA CTCGGGCCAG      1800
      AACGCCTGGC TGTCGTACCA GCTGCTCAAG GCCACGGAGC CGGGGCTGTT CGGCGTGTGG    1860
      GCGCACAATG GCGAGGTGCG CACCGCCAGG CTGCTGAGGG AGCGCGACGC TGCCAAGCAG    1920
      AGGCTGGTGG TGCTGGTCAA GGACAATGGC GAGCCTCGGC GCTCGGCCAC CGCCACGCTG    1980
      CACGTGCTCC TGGTGGACGG CTTCTCCCAG CCCTACCTGC TGCTCCCCGGA GGCGGCACCG    2040
      GCCCAGGCCC AGGCCGACTT GCTCACCGTC TACCTGGTGG TGGCGTTGGC CTCGGTGTCT    2100
      TCGCTCTTCC TCTTCTCGGT GCTCCTGTTC GTGGCGGTGG GGCTGTGCAG GAGGAGCAGG    2160
      GCGGCCTCGG TGGGTCGCTG CTCGGTGCCC GAGGGCCCCT TTCCAGGGCA GATGGTGGAC    2220
      GTGAGCGGCA CCGGGACCCT GTCCCAGAGC TACCAGTACG AGGTGTGTCT GACTGGAGGC    2280
      TCCGGGACAA ATGAGTTCAA GTTCCTGAAG CCAATTATCC CCAACTTCGT TGCTCAGGGT    2340
      GCAGAGAGGG TTAGCGAGGC AAATCCCAGT TTCAGGAAGA GCTTTGAATT CACTTAA      2397
```

Seq ID NO: 288 <u>Protein sequence</u>
Protein Accession #: AAD43756.1

```
       1         11        21        31        41        51
       |         |         |         |         |         |
      MEAGEGKERV PKQRQVLIFF VLLGIAQASC QPRHYSVAEE TESGSFVANL LKDLGLEIGE    60
      LAVRGARVVS KGKKMHLQPD RQTGDLLLNE KLDREELCGP TEPCVLPPQV LLENPLQPFQ    120
      AELRIRDVND HSPVFLDKEI LLKIPESITP GTTFLIERAQ DLDVGTNSLQ NYTISPNFHF    180
      HLNLQDSLDG IILPQLVLNR ALDREEQPEI RLTLTALDGG SPPRSGTALV RIEVVDINDN    240
      VPEFAKLLYE VQIPEDSPVG SQVAIVSARD LDIGTNGEIS YAFSQASEDI EKTFRLSAKS    300
      GELLLRQKLD FESIQTYTVN IQATDGGGLS GTCVVFVQVM DLNDNPPELT MSTLINQIPE    360
      NLQDTLIAVF SVSDPDSGDN GRMVCSIQDD LPFFLKPSVE NFYTLVISTA LDRETRSEYN    420
      ITITVTDPGT PRLKTEHNIT VLVSDVNDNA PAFTQTSYTL FVRENNSPAL HIGSVSATDR    480
      DSGTNAQVTY SLLPPQDPHL PLASLVSINA DNGHLFALQS LDYEALQAFE FRVGAADRGS    540
      PALSSEALVR VLVLDANDNS PFVLYPLQNG SAPCTELVPR AAEPGYLVTK VVAVDGDSGQ    600
      NAWLSYQLLK ATEPGLFGVW AHNGEVRTAR LLRERDAAKQ RLVVLVKDNG EPPRSATATL    660
      HVLLVDGFSQ PYLLLPEAAP AQRAQADLLTV YLVVALASVS SLFLFSVLLF VAVRLCRRSR    720
      AASVGRCSVP EGPFPGQMVD VSGTGTLSQS YQYEVCLTGG SGTNEFKFLK PIIPNFVAQG    780
      AERVSEANPS FRKSFEFT
```

Seq ID NO: 289 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_018674.1
Coding sequence: 390..2009

```
       1         11        21        31        41        51
       |         |         |         |         |         |
```

```
         CGGAGCACAT GCTGAGCGGA GCGGCTGGGG CTGCGCGGCG TGGCGGAGCA GCGCTCGCTC    60
         CCTCGCTCAC TCGCTCGCTC GCAGGGACAC ACGCAGGGGC TGACAGCTGT GCTGGTGCTG   120
         ATAAGGGAAG CCACAAGGAG ACGATCGAGG AGAGAGACAA GCGGCAGCAG AGGCAGCAGC   180
5        GGCAGAGGCA GCACCAGGGC TGCGGAGCTG CTGGGAGTGG GAGTGACTCC CCCACCTCGG   240
         GCCCCCACCC TGTCCCTGTC CTCTTCCCGC TTGCCCTGAG TTTAGAAGAG CAGCCGCTGC   300
         CACCACTGCC ACTCGGGAGG GCACCAGGGC TGCTGGCTAG GGAGGGACAG GGCAGGGAGG   360
         CTCTGGCCAG TCCCAGCAGC CGGGGACAGA TGCCGATCGA GATTGTGTGC AAAATCAAAT   420
         TTGCTGAGGA GGATGCGAAA CCCAAGGAGA AGGAGGCAGG GGATGAGCAG AGCCTCCTCG   480
10       GGGCTGTTGC CCCTGGAGCA GCCCCCCGAG ACCTGGCCAC CTTTGCCAGC ACCAGCACCC   540
         TGCATGGACT GGGCCGGGCC TGTGGCCCAG GCCCCCACGG ACTGCGCAGA ACCCTGTGGG   600
         CACTGGCCCT ACTCACCTCG CTGGCTGCCT TCCTGTACCA GGCGGCTGCC CTGGCCCGGG   660
         GCTACCTGAC CCGGCCTCAC CTGGTGGCAA TGGACCCCGC TGCCCCAGCC CCAGTGGCGG   720
         GCTTCCCGGC TGTCACCCTC TGCAATATCA ACCGCTTCCG GCATTCGGCA CTCAGCGATG   780
15       CCGACATCTT CCACCTGGCC AATCTGACAG GGCTGCCCCC CAAAGACCGG GATGGGCACC   840
         GTGCGGCTGG CCTGCCGCTAC CCAGAGCCTG ACATGGTAGA CATCCTCAAC CGCACTGGCC   900
         ACCAGCTCGC CGACATGCTT AAGAGCTGCA ACTTCAGTGG GCATCACTGC TCCGCCAGCA   960
         ACTTCTCTGT CGGTCTATACT CGCTATGGGA AGTGTTACAC CTTCAACGCG GACCCGCGGA  1020
         GCTCGCTGCC CAGCCGGGCA GGGGGCATGG GCAGTGGGCCT GGAGATCATG CTGGACATCC  1080
20       AGCAGGAGGA GTACCTGCCC ATCTGGAGGG AGACAAATGA GACGTCGTTT GAGGCAGGTA  1140
         TTCGGGTGCA GATCCACAGC CAGGAGGAGC CGCCCTACAT CCACCAGCTG GGGTTCGGGG  1200
         TGTCCCCAGG CTTCCAGACC TTTGTGTCCT GCCAGGAACA GCGGCTGACC TACCTGCCCC  1260
         AGCCCTGGGG CAACTGCCGC GCAGAGAGTG AGCTCAGGGA GCCTGAGCTT CAGGGCTACT  1320
         CGGCCTACAG TGTGTCTGCC TGCCGGCTGC GCTGTGAAAA GGAGGCCGTG CTTCAGCGCT  1380
25       GCCACTGCCG GATGGTGCAC ATGCCAGGCA ATGAGACCAT CTGCCCACCA AATATCTACA  1440
         TCGAGTGTGC AGACCACACA CTGGACTCCC TGGGTGGGGG CCCTGAGGGC CCGTGCTTCT  1500
         GCCCCACCCC CTGCAACCTG ACACGCTATG GGAAAGAGAT CTCCATGGTC AGGATCCCCA  1560
         ACAGGGGCTC AGCCCGGTAC CTGGCGAGGA AGTACAACCG CAACGAGACC TACATACGGG  1620
         AGAACTTCCT GGTCCTAGAT GTCTTCTTTG AGGCCCTGAC CTCTGAAGCC ATGGAGCAGC  1680
30       GAGCAGCCTA TGGCCTGTCA GCCCTGCTGG GAGACCTCGG GGGACAGATG GGCCTGTTCA  1740
         TTGGGGCCAG CATCCTCACG TTGCTGGAGA TCCTCGACTA CATCTATGAG GTGTCCTGGG  1800
         ATCGACTGAA GCGGGTATGG AGGCGTCCCA AGACCCCCCT GCGGACCTCC ACTGGGGGCA  1860
         TCTCCACTTT GGGGCTTCAG GAGCTGAAGG AACAGAGTCC CTGCCCGAGC CTGGGCCCAG  1920
         CGGAGGGTGG GGGGGTCAGC AGTCTGCTCC CCAATCACCA CCACCCCCAC GGTCCCCCAG  1980
35       GAGGTCTCTT TGAAGATTTT GCTTGCTAGG ACGGTGCTGT GACTGAAAGG ACCCAGGAGT  2040
         CTGGGACCCC TCCTGGGATC CCCAGCACAT TCTCCTGCTC CTGGGAGAGG CCTGGGGGCG  2100
         GTGCTCACTG GGAGGGCCAG GACTCAGTTC CTGCTCTCAT CCTCCCCTGC CCTGATGTCA  2160
         GCTGCTTTGC ACAAAGGTCC TTCTTGTCCA CACCCCTTAT CCCCAGGCTG GTGCCCCGGG  2220
         AGGGCTGGAG ACCAGGCCAT GGGCCCTCAC GGAGAGGAAG GGAAGGAAGG AGAGGGAGGG  2280
40       GGAGGATAGA GCCCATCCCA GCCGGGGAGG GGGAGCCCTC TGTACATTTG TAAATATTTA  2340
         GGGAAAGCCG GGTGGGGGGA GGGGATACAG ATGTAGAAGG TGGGTAGGGC TACAGGGGTG  2400
         GGTGATTTAG GGACAGCCAG GGTCCCAGCC CCAATGTCAG CAGGATAGGG AGAGCCCCAG  2460
         GACTCAGGAG TGCTGGGGCTG GTCCTACTTC CTGCCCCTCT CCAGGCCCAG CTCCCCTCTT  2520
         GGCAGGGGGA GAGGATGGCC CAGCAGGCCT GGCCCAGCTC CCAGTTCCCC CTGCACCAGC  2580
45       CCCACCCCTA GAGTCCCTTC TATAGGGAGG GGGCAGGAGA CCTTCCAGAC TTCGGCTGAG  2640
         CTTGGAGGGT GGGAAGGGAG CCTTCTCAGT CCTCTCTCCC TCCAGTCTGA TTTTATAAAG  2700
         TGCTGACGAG
```

Seq ID NO: 290 Protein sequence
Protein Accession #: NP_061144.1

```
50
     1          11         21         31         41         51
     |          |          |          |          |          |
     MPIEIVCKIK FAEEDAKPKE KEAGDEQSLL GAVAPGAAPR DLATFASTST LHGLGRACGP    60
     GPHGLRRTLW ALALLTSLAA FLYQAAGLAR GYLTRPHLVA MDPAAPAPVA GFPAVTLCNI   120
55   NRFRHSALSD ADIFHLANLT GLPPKDRDGH RAAGLRYPEP DMVDILNRTG HQLADMLKSC   180
     NFSGHHCSAS NFSVVYTRYG KCYTFNADPR SSLPSRAGGM GSGLEIMLDI QQEEYLPIWR   240
     EINETSFEAG IRVQIHSQEE PPYIHQLGFG VSPGFQTFVS CQEQRLTYLP QFWGNCRAES   300
     ELREPELQGY SAYSVSACRL RCEKEAVLQR CHCRMVHMPG NETICPPNIY IECADHTLDS   360
     LGGGPEGPCF CPTPCNLTRY GKEISMVRIP NRGSARYLAR KYNRNETYIR ENFLVLDVFF   420
60   EALTSEAMEQ RAAYGLSALL GDLGGQMGLF IGASILTLLE ILDYIYEVSW DRLKRVWRRP   480
     KTPLRTSTGG ISTLGLQELK EQSPCPSLGR AEGGGVSSLL PNHHHPHGPP GGLFEDFAC
```

Seq ID NO: 291 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 62..895

```
65
     1          11         21         31         41         51
     |          |          |          |          |          |
     CACTGCTCTG AGAATTTGTG AGCAGCCCCT AACAGGCTGT TACTTCACTA CAACTGACGA    60
     TATGATCATC TTAATTTACT TATTTCTCTT GCTATGGGAA GACACTCAAG GATGGGGATT   120
     CAAGGATGGA ATTTTTCATA ACTCCATATG GCTTGAACGA GCAGCCGGTG TGTACCACAG   180
70   AGAAGCACGG TCTGGCAAAT ACAAGCTCAC CTACGCAGAA GCTAAGGCGG TGTGTGAATT   240
     TGAAGGCGGC CATCTCGCAA CTTACAAGCA GCTAGAGGCA GCCAGAAAAA TTGGATTTCA   300
     TGTCTGTGCT GCTGGATGGA TGGCTAAGGG CAGAGTTGGA TACCCCATTG TGAAGCCAGG   360
     GCCCAACTGT GGATTTGAA AAACTGGCAT TATTGATTAT GGAATCCGTC TCAATAGGAG   420
     TGAAAGATGG GATGCCTATT GCTACAACCC ACACGCAAAG GAGTGTGTG GCGTCTTTAC   480
75   AGATCCAAAG CAAATTTTTA AATCTCCAGG CTTCCCAAAT GAGTACGAAG ATAACCAAAT   540
     CTGCTACTGG CACATTAGAC TCAAGTATGG TCAGCGTATT CACCTGAGTT TTTTAGATTT   600
     TGACCTTGAA GATGACCCAG GTTGCTTGGC TGATTATGTT GAAATATATG ACAGTTACGA   660
     TGATGTCCAT GGCTTTGTGG GAAGATACTG TGGAGATGAG CTTCCAGATG ACATCCATCAG   720
     TACAGGAAAT GTCATGACCT TGAAGTTTCT AAGTGATGCT TCAGTGACAG CTGGAGGTTT   780
80   CCAAATCAAA TATGTTGCAA TGGATCCCTGT ATCCAAATCC AGTCAAGGAA AAAATACAAG   840
     TACTACTTCT ACTGGAAATA AAAACTTTTT AGCTGGAAGA TTTAGCCACT TATAAAAAAA   900
     AAAAAAAGGA TGATCAAAAC ACACAGTGTT TATGTTGGAA TCTTTTGGAA CTCCTTTGAT   960
```

```
     CTCACTGTTA TTATTAACAT TTATTTATTA TTTTTCTAAA TGTGAAAGCA ATACATAATT   1020
     TAGGGAAAAT TGGAAAATAT AGGAAACTTT AAACGAGAAA ATGAAACCTC TCATAATCCC   1080
     ACTGCATAGA AATAACAAGC GTTAACATTT TCATATTTTT TTCTTTCAGT CATTTTTCTA   1140
     TTTGTGGTAT ATGTATATAT GTACCTATAT GTATTTGCAT TTGAAATTTT GGAATCCTGC   1200
  5  TCTATGTACA GTTTTGTATT ATACTTTTTA AATCTTGAAC TTTATAAACA TTTTCTGAAA   1260
     TCATTGATTA TTCTACAAAA ACATGATTTT AAACAGCTGT AAAATATTCT ATGATATGAA   1320
     TGTTTTATGC ATTATTTAAG CCTGTCTCTA TTGTTGGAAT TTCAGGTCAT TTTCATAAAT   1380
     ATTGTTGCAA TAAATATCCT TGAACACACA AAAAAAAAAA AA
```

10 Seq ID NO: 292 Protein sequence
Protein Accession #: Eos sequence

```
     1          11         21         31         41         51
     |          |          |          |          |          |
  15 MIILIYLFLL LWEDTQGWGF KDGIFHNSIW LERAAGVYHR EARSGKYKLT YAEAKAVCEF    60
     EGGHLATYKQ LEAARKIGFH VCAAGWMAKG RVGYPIVKPG PNCGFGKTGI IDYGIRLNRS   120
     ERWDAYCYNP HAKECGGVFT DPKQIFKSPG FPNEYEDNQI CYWHIRLKYG QRIHLSFLDF   180
     DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELPDDIIS TGNVMTLKFL SDASVTAGGF   240
  20 QIKYVAMDFV SKSSQGKNTS TTSTGNKNFL AGRFSHL
```

Seq ID NO: 293 DNA sequence
Nucleic Acid Accession #: NM_007115.1
Coding sequence: 69..902

```
     1          11         21         31         41         51
     |          |          |          |          |          |
  25 GAATTCGCAC TGCTCTGAGA ATTTGTGAGC AGCCCCTAAC AGGCTGTTAC TTCACTACAA    60
     CTGACGATAT GATCATCTTA ATTTACTTAT TTCTCTTGCT ATGGGAAGAC ACTCAAGGAT   120
     GGGGATTCAA GGATGGAATT TTTCATAACT CCATATGGCT TGAACGAGCA GCCGGTGTGT   180
  30 ACCACAGAGA AGCACGGTCT GGCAAATACA AGCTCACCTA CGCAGAAGCT AAGGCGGTGT   240
     GTGAATTTGA AGGCGGCCAT CTCGCAACTT ACAAGCAGCT AGAGGCAGCC AGAAAAATTG   300
     GATTTCATGT CTGTGCTGCT GGATGGATGG CTAAGGGCAG AGTTGGATAC CCCATTGTGA   360
     AGCCAGGGCC CAACTGATGA TTTGGAAAAA CTGGCATTAT TGATTATGGA ATCCGTCTCA   420
     ATAGGAGTGA AAGATGGGAT GCCTATTGCT ACAACCCACA GGCAAAGGAG TGTGGTGGCG   480
  35 TCTTTACAGA TCCAAAGCGA ATTTTTAAAT CTCCAGGCTT CCCAAATGAG TACGAAGATA   540
     ACCAAATCTG CTACTGGCAC ATTAGACTCA AGTATGGTCA GCGTATTCAC CTGAGTTTTT   600
     TAGATTTTGA CCTTGAAGAT GACCCAGGTT GCTTGGCTGA TTATGTTGAA ATATATGACA   660
     GTTACGATGA TGTCCATGGC TTTGTGGGAA GATACTGTGG AGATGAGCTT CCAGATGACA   720
     TCATCAGTAC AGGAAATGTC ATGACCTTGA AGTTTCTAAG TGATGCTTCA GTGACAGCTG   780
  40 GAGGTTTCCA AATCAAATAT GTTGCAATGG ATCCTGTATC CAAATCCAGT CAAGGAAAAA   840
     ATACAAGTAC TACTTCTACT GGAAATAAAA ACTTTTTAGC TGGAAGATTT AGCCACTTAT   900
     AAAAAAAAAA AAGGATGATC AAAACACACA GTGTTTATGT TGGAATCTTT TGGAACTCCT   960
     TTGATCTCAC TGTTATTATT AACATTTATT TATTATTTTT CTAAATGTGA AAGAAATACA   1020
     TAATTTAGGG AAAAATTGGA AATATAGGAA ACTTTAAACG AGAAAATGAA ACCTCTCATA   1080
  45 ATCCCACTGC ATAGAAATAA CAAGCGTTAA CATTTTCATA TTTTTTTCTT TCAGTCATTT   1140
     TTGTATTTGT GGTATATGTA TATATGTACC TATATGTATT TGCATTTGAA ATTTTGGAAT   1200
     CCTGCTCTAT GTACAGTTTT GTATTATACT TTTTAAATCT TGAACTTTAT GAACATTTTC   1260
     TGAAATCATT GATTATTCTA CAAAAACATG ATTTTAAACA GCTGTAAAAT ATTCTATGAT   1320
     ATGAATGTTT TATGCATTAT TTAAGCCTGT CTCTATTGTT GGAATTTCAG GTCATTTTCA   1380
  50 TAAATATTGT TGCAATAAAT ATCCTTCGGA ATTC
```

Seq ID NO: 294 Protein sequence
Protein Accession #: NP_009046.1

```
     1          11         21         31         41         51
     |          |          |          |          |          |
  55 MIILIYLFLL LWEDTQGWGF KDGIFHNSIW LERAAGVYHR EARSGKYKLT YAEAKAVCEF    60
     EGGHLATYKQ LEAARKIGFH VCAAGWMAKG RVGYPIVKPG PNXXFGKTGI IDYGIRLNRS   120
     ERWDAYCYNP HAKECGGVFT DPKRIFKSPG FPNEYEDNQI CYWHIRLKYG QRIHLSFLDF   180
  60 DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELPDDIIS TGNVMTLKFL SDASVTAGGF   240
     QIKYVAMDFV SKSSQGKNTS TTSTGNKNFL AGRFSHL
```

65 Seq ID NO: 295 DNA sequence
Nucleic Acid Accession #: NM_001218.2
Coding sequence: 116..1180

```
     1          11         21         31         41         51
     |          |          |          |          |          |
  70 GTACTCGCCA CGGCACCCAG GCTGCGCGCA CGCGGTCCCG GTGTGCAGCT GGAGAGCGAG    60
     CGGCCACCGG GAGCCCCCGG CACAGCCCGC GCCCGCCCCG CAGGAGCCCG CGAAGATGCC   120
     CCGGCGCAGC CTGCACGCGG CGGCCGTGCT CCTGCTGGTG ATCTTAAAGG AACAGCCTTC   180
     CAGCCCGGCC CCAGTGAACG GTTCCAAGTG GACTTATTTT GGTCCTGATG GGGAGAATAG   240
  75 CTGGTCCAAG AAGTACCCGT CGTGTGGGGG CCTGCTGCAG TCCCCCATAG ACCTGCACAG   300
     TGACATCCTC CAGTATGACG CCAGCCTCAC GCCCCTCGAG TTCCAAGGCT ACAATCTGTC   360
     TGCCAACAAG CAGTTTCTCC TGACCAACAA TGGCCATTCA GTGAAGCTGA ACCTGCCCTC   420
     GGACATGCAC ATCCAGGGCC TCCAGTCTCG CTACAGTGCC ACGCAGCTGC ACCTGCACTG   480
     GGGGAACCGG AATGACCCGC ACGGCTCTGA GCACACCGTC ACTTCGCCGC   540
  80 CGAGCTGCAC ATTGTCCATT ATAACTCAGA CCTTTATCCT GACGCCAGCA CTGCCCAGCA   600
     CAAGTCAGAA GGCCTCGCTG TCCTGGCTGT TCTCATTGAG ATGGGCTCCT TCAATCCGTC   660
     CTATGACAAG ATCTTCAGTC ACCTTCAACA TGTAAAGTAC AAAGGCCAGG AAGCATTCGT   720
     CCCGGGATTC AACATTGAAG AGCTGCTTCC GGAGAGGACC GCTGAATATT ACCGCTACCG   780
     GGGGTCCCTG ACCACACCCC CTTGCAACCC CACTGTGCTC TGGACAGTTT TCCGAAACCC   840
     CGTGCAAATT TCCCAGGAGC AGCTGCTGGC TTTGGAGACA GCCCTGTACT GCACACACAT   900
```

```
GGACGACCCT TCCCCCAGAG AAATGATCAA CAACTTCCGG CAGGTCCAGA AGTTCGATGA   960
GAGGCTGGTA TACACCTCCT TCTCCCAAGT GCAAGTCTGT ACTGCGGCAG GACTGAGTCT  1020
GGGCATCATC CTCTCACTGG CCCTGGCTGG CATTCTTGGC ATCTGTATTG TGGTGGTGGT  1080
GTCCATTTGG CTTTTCAGAA GGAAGAGTAT CAAAAAAGGT GATAACAAGG GAGTCATTTA  1140
CAAGCCAGCC ACCAAGATGG AGACTGAGGC CCACGCTTGA GGTCCCCGGA GCTCCCGGGC  1200
ACATCCAGGA AGGACCTTGC TTTGGACCCT ACACACTTCG GCTCTCTGGA CACTTGCGAC  1260
ACCTCAAGGT GTTCTCTGTA GCTCAATCTG CAAACATGCC AGGCCTCAGG GATCCTCTGC  1320
TGGGTGCCTC CTTGCCTTGG GACCATGGCC ACCCCAGAGC CATCCGATCG ATGGATGGGA  1380
TGCACTCTCA GACCAAGCAG CAGGAATTCA AAGCTGCTTG CTGTAACTGT GTGAAGATTGT  1440
GAAGTGGTCT GAATTCTGGA ATCACAAACC AAGCCATGCT GGTGGGCCAT TAATGGTTGG  1500
AAAACACTTT CATCCGGGGC TTTGCCAGAG CGTGCTTTCA AGTGTCCTGG AAAGTCTGAT  1560
GCTTCTCCAA GCTTTCAGAC AAGAATGTGC ACTCTCTGCT TAGGTTTTGC TTGGGAAACT  1620
CAACTTCTTT CCTCTGGAGA CGGGGCATCT CCCTCTGATT TCCTTCTGCT ATGACAAAAC  1680
CTTTAATCTG CACCTTACAA CTCGGGGACA AATGGGGACA GGAAGGATCA AGTTGTAGAG  1740
AGAAAAAAGA AAACAAGAGA TATACATTGT GATATATTAG GACACTTTC ACAGTCCTGT  1800
CCTCTGGATC ACAGACACTG CACAGACCTT AGGGAATGGC AGGTTCAAGT TCCACTTCTT  1860
GGTGGGGATG AGAAGGGAGA GAGAGCTAGA GGGACAAAGA GAATGAGAAG ACATGGATGA  1920
TCTGGGAGAG TCTCACTTTG GAATCAGAAT TGGAATCACA TTCTGTTTAT CAAGCCATAA  1980
TGTAAGGACA GAATAATACA ATATTAAGTC CAAATCCAAC CTCCTGTCAG TGGAGCAGTT  2040
ATGTTTTATA CTCTACAGAT TTTACAAATA ATGAGGCTGT TCCTTGAAAA TGTGTTGTTG  2100
CTGTGTCCTG GAGGAGACAT GAGTTCCGAG ATGACCCAAT CTGCCTTTGA ATCTGGAGGA  2160
AATAGGCAGA AACAAAATGA CTGTAGAACT TATTCTCTGT AGGCCAAATT TCATTTCAGC  2220
CACTTCTGCA GGATCCCTAC TGCCAACCTG GAATGGAGAC TTTTATCTAC TTCTCTCTCT  2280
CTGAAGATGT CAAATCGTGG TTTAGATCAA ATATATTTCA AGCTATAAAA GCAGGAGGTT  2340
ATCTGTGCAG GGGGCTGGCA TCATGTATTT AGGGGCAAGT AATAATGGAA TGCTACTAAG  2400
ATACTCCATA TTCTTCCCCG AATCACACAG ACAGTTTCTG ACAGGCGCAA CTCCTCCATT  2460
TTCCTCCCGC AGGTGAGAAC CCTGTCGGAGA TGAGTCAGTG CCATGACTGA GAAGGAACCG  2520
ACCCCTAGTT GAGAGCACCT TGCAGTTCCC CGAGAACTTT CTGATTCACA GTCTCATTTT  2580
GACAGCATGA AATGTCCTCT TGAAGCATAG CTTTTTAAAT ATCTTTTTCC TTCTACTCCT  2640
CCCTCTGACT CTAAGAATTC TCTCTTCTGG AATCGCTTGA ACCCAGGAGG CGGAGGTTGC  2700
AGTAAGCCAA GGTCGTGCCA CTGCACTCTA GCCTGGGTGA CAGAGCGAGA CTCCATCTCA  2760
AAAAAAAAAA AAAA
```

Seq ID NO: 296 Protein sequence
Protein Accession #: NP_001209.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MPRRSLHAAA VLLLVILKEQ PSSPAPVNGS KWTYPGPDGE NSWSKKYPSC GGLLQSPIDL   60
HSDILQYDAS LTPLPFQGYN LSANKQFLLT KNGHSVKLNL PSDMHIQGLQ SRYSATQLHL  120
HWGNPNDPHG SEHTVSGQHF AAELHIVHYN SDLYPDASTA SNKSEGLAVL AVLIEMGSFN  180
PSYDKIFSHL QEVKYKGQEA FVPGFNIEEL LPERTAEYYR YRGSLTTPPC NPTVLWTVPR  240
NPVQISQEQL LALETALYCT HMDDPSPREM INNFRQVQKF DERLVYTSFS QVQVCTAAGL  300
SLGIILSLAL AGILGICIVV VVSIWLFRRK SIKKGDNKGV IYKPATKMET EAHA
```

Seq ID NO: 297 DNA sequence
Nucleic Acid Accession #: NM_006632.1
Coding sequence: 377..1582

```
1          11         21         31         41         51
|          |          |          |          |          |
ACGCGTCCGC CCACGCGTCC GCCCACGCGT CCGGTCGGGG CCAGAGCGCA GGTGTACCTG   60
GCGGCCGTGC TGGAGCACCT GACCGCCGAG ATCCTGGACG TGGCTGGCAA CCCCGGCCGC  120
GACAAGAAGA CCCGCATCAT CCTGCGCCAC CTGTAGCTGG CCATTCGCAA CGGCGAGGAG  180
CTTAACAAGC TGCTGGGCGA AGTCACCATC GCGCAGGGCG GTGTCCTGCC CAACATTCAG  240
GGCGTGCTTC TGCCCCAGAA GACCAAGAGC CACCACAAGG CCAAGGGTGA AAACCATTCA  300
CTAGGAGAGG AGAAACACAA TGGCCACCAA GACAGAGTTG AGTCCCACAG CAAGGGAGAG  360
CAAGAACGCA CAAGATATGC AAGTGGATGA GACACTGATC CCCAGGAAAG GTCCAAGTTT  420
ATGTTCTGCT CGCTATGGAA TAGCCCTCGT CTTACATTTC TGCAATTTCA CAACGATAGC  480
ACAAAATGTC ATCATGAACA TCACCATGGT AGCCATGGTC AACAGCACAA GCCCTCAATC  540
CCAGCTCAAT GATTCCTCTG AGGTGCTGCC TGTTGACTCA TTTGGTGGCC TAAGTAAAGC  600
CCCAAGAGT CTTCCTGCAA AGTCCTCAAT ACTTGGGGGT CAGTTTGCAA TTTGGGAAAA  660
GTGGGGCCCT CCACAAGAAC GAAGCAGACT CTGCAGCATT GCTTTATCAG GAATGTTACT  720
GGGATGCTTT ACTGCCATCC TCATAGGTGG CTTCATTAGT GAAACCCTTG GGTGGCCCTT  780
TGTCTTCTAT ATCTTTGGAG GTGTTGGCTG TGTCTGCTGC CTTCTCTGGT TTGTTGTGAT  840
TTATGATGAC CCCTTTTCCT ATCCATGGAT AAGCACCTCA GAAAAAGAAT ACATCATATC  900
CTCCTTGAAA CAACAGGTCG GGTCTTCTAA GCAGCCTCTT CCCATCAAAG CTATGCTCAG  960
ATCTCTACCC ATTTGGTCCA TATGTTTAGG CTGTTTCAGC CATCAATGGT TAGTTAGCAC  1020
AATGGTTGTA TACATACCAA CTTACATCAG CTCTGTGTAC CATGTTAACA TCAGAGACAA  1080
TGGACTTCTA TCTGCCCTTC CTTTTATTGT TGCCTGGGTC ATAGGCATGG TGGGAGGCTA  1140
TCTGGCAGAT TTCCTTCTAA CCAAAAAGTT TAGACTCATC ACTGTGAGGA AAATTGCCAC  1200
AATTTTAGGA AGTCTCCCCT CTTCAGCACT CATTGTGTCT GTGCCTTACC TCAATTCCGG  1260
CTATATCACA GCAACTGCCT TGCTGACGCT CTCTTGCGGA TTAAGCACAT TGTGTCAGTC  1320
AGGGATTTAT ATCAATGTCT TAGATATTGC TCCAAGGTAT TCCAGTTTTC TCATGGGAGC  1380
ATCAAGAGGA TTTTCGAGCA TAGCACCTGT CATTGTACCC ACTGTCAGCG GATTTCTTCT  1440
TAGTCAGGAC CCTGAGTTTG GGTGGAGGAA TGTCTTCTTC TTGCTGTTTG CCGTTAACCT  1500
GTTAGGACTA CTCTTCTACC TCATATTTGG AGAAGCAGAT GTCCAAGAAT GGGCTAAAGA  1560
GAGAAAACTC ACTCGTTTAT GAAGTTATCC CACCTTGGAT GGAAAAGTCA TTAGGCACCG  1620
TATTGCATAA AATAGAAGGC TTCCGTGATG AAAATACCAG TGAAAAGATT TTTTTTCCT  1680
GTGGCTCTTT TCAATTATGA GATCAGTTCA TTATTTATT CAGACTTTTT TTTGAGAGAA  1740
ATGTAAGATG AATAAAAATT CAAATAAAAT GATAACTAAG AAAAAAAAAA AAAAA
```

Seq ID NO: 298 Protein sequence
Protein Accession #: NP_006623.1

```
       1          11         21         31         41         51
       |          |          |          |          |          |
     MQVDETLIPR KGPSLCSARY GIALVLHFCN FTTIAQNVIM NITMVAMVNS TSPQSQLNDS    60
5    SEVLPVDSFG GLSKAPKSLP AKSSILGGQF AIWEKWGPPQ ERSRLCSIAL SGMLLGCFTA   120
     ILIGGPISET LGWPFVFYIF GGVGCVCCLL WFVVIYDDPF SYPWISTSEK EYIISSLKQQ   180
     VGSSKQPLPI KAMLRSLPIW SICLGCFSHQ WLVSTMVVYI PTYISSVYHV NIRDNGLLSA   240
     LPFIVAWVIG MVGGYLADFL LTKKFRLITV RKIATILGSL PSSALIVSLP YLNSGYITAT   300
     ALLTLSCGLS TLCQSGIYIN VLDIAPRYSS FLMGASRGFS SIAPVIVPTV SGFLLSQDPE   360
10   FGWRDVFFLL FAVNLLGLLF YLIFGEADVQ EWAKERKLTR L
```

Seq ID NO: 299 DNA sequence
Nucleic Acid Accession #: NM_003058.1
Coding sequence: 145..1812

```
       1          11         21         31         41         51
       |          |          |          |          |          |
     GGCCCTGCCC TGAAGGCTGG TCACTTGCAG AGGTAAACTC CCCTCTTTGA CTTCTGGCCA    60
     GGGTTGTGC  TGAGCTGGCT GCAGCCGCTC TCAGCCTCGC TCCGGGCACG TCGGGCAGCC   120
20   TCGGGCCCTC CTGCCTGCAG GATCATGCCC ACCACCGTGG ACGATGTCCT GGAGCATGGA   180
     GGGGAGTTTC ACTTTTCCA  GAAGCAAATG TTTTTCCTCT TGGCTCTGCT CTCGGCTACC   240
     TTCGCGCCCA TCTACGTGGG CATCGTCTTC CTGGGCTTCA CCCCTGACCA CCGCTGCCGG   300
     AGCCCGGAG  TGGCCGAGCT GAGTCTCGCC TGCGGCTGGA GTCCTGCAGA GGAACTGAAC   360
     TACACGGTGC CGGGCCCAGG ACCTGCGGGC GAAGCCTCCC CAAGACAGTG TAGGCGCTAC   420
25   GAGGTGGACT GGAACCAGAG CACCTTTGAC TGCGTGGACC CCCTGGCCAG CCTGGACACC   480
     AACAGGAGCC GCCTGCCACT GGGCCCCTGC CGGGACGGCT GGGTGTACGA GACGCCTGGC   540
     TCGTCCATCG TCACCGAGTT TAACCTGGTA TGTGCCAACT CCTGGATGTT GGACCTATTC   600
     CAGTCATCAG TGAATGTAGG ATTCTTTATT GGCTCTATGA GTATCGGCTA CATAGCCAGC   660
     AGGTTTGGCC GTAAGCTCTG CCTCCTAACT ACAGTCCTCA TAAATGCTGC AGCTGGAGTT   720
30   CTCATGGCCA TTTCCCCAAC CTATACGTGG ATGTTAATTT TTCGCTTAAT CCAAGGACTG   780
     GTCAGCAAAG CAGGCTGGTT AATAGGCTAC ATCCTGATTA CAGAATTTGT TGGGCGGAGA   840
     TATCGGAGAA CAGTGGGGAT TTTTTACCAA GTTGCCTATA CAGTTGGGCT CCTGGTGCTA   900
     GCTGGGGTGG CTACGGCACT TCCTCACTGG AGGTGGTTGC AGTTCACAGT TGCTCTGCCC   960
     AACTTCTTCT TCTTGCTCTA TTACTGGTGC ATACCTGAGT CTCCCAGGTG GCTGATCTCC  1020
35   CAGAATAAGA ATGCTGAAGC CATGAGAATC ATTAAGCACA TCGCAAAGAA AAATGGAAAA  1080
     TCTCTACCCG CCTCCCTTCA GCGGCCTGAGA CTTGAAGAGG AAACTGGCAA GAAATTGAAC  1140
     CCTTCATTTC TTGACTTGGT CAGAACTCCT CAGATAAGGA AACATACTAT GATATTGATG  1200
     TACAACTGGT TCACGAGCTC TGTGCTCTAC CAGGGCCTCA TCATGCACAT GGGGCCTTGCA  1260
     GGTGACAATA TCTACCTGGA TTTCTTCTAC TCTGCCCTGG TTGAATTCCC AGCTGCCTTC  1320
40   ATGATCATCC TCACCATCGA CCGGCATCGGA CGCCGTTACC CTTGGGCTGC ATCAAATATG  1380
     GTTGCAGGGG CAGCCTGTCT GGCCTCAGTT TTTATACCTG GTGATCTACA ATGGCTAAAA  1440
     ATTATTATCT CATGCTTGGG AAGAATGGGG ATCACAATGG CCTATGAGAT AGTCTGCCTG  1500
     GTCAATGCTG AGCTGTACCC CACATTCATT AGGAATCTTG GCGTCCACAT CTGTTCCTCA  1560
     ATGTGTGACA TTGGTGGCAT CATCACGCCA TTCCTGGTCT ACCGGCTCAC TAACATCTGG  1620
45   CTTGAGCTCC CGCTGATGGT TTTCGGCGTA CTTGGCTTGG TTGCTGGAGG TCTGGTGCTG  1680
     TTGCTTCCAG AAACTAAAGG GAAAGCTTTG CCTGAGACCA TCGAGGAAGC CGAAAATATG  1740
     CAAAGACCAA GAAAAAATAA AGAAAAGATG ATTTACCTCC AAGTTCAGAA ACTAGACATT  1800
     CCATTGAACT AAGAAGAGAG ACCGTTGCTG CTGTCATGAC CTAGCTTTGA TGGCAGCAAG  1860
     ACCAAAAGTA GAAATCCCTG CACTCATCAC AAAGCCCATA CAACTCAACC AAACTTACCC  1920
50   CTGAGCCCTA TCAACCTAGG TCTACAGCCA GTGGAGTCTA TTGTACACTG TGGAAAAATA  1980
     CCCATGGGAC CAGATCCTGC CAAATTCTTC CAGCTCACTT TATTCTCAGC ATTCCTAGGA  2040
     CATTGGACAT TGGTTTTCTG GAGGGTTTTT TTTCCGATCT TTGTATTTTT TTAAATTTGA  2100
     TTCTTTTCTT TGCAATGCTA GCAACCAGAA TACATAGGGG AACTGTGGGC TAGGCAAAA   2160
     AAATAGAAAA AGTGTGAAAA ACAGTAAAGT TGGGAGAGGA GCATCTATTT TCTTAAAGAA  2220
55   ATAAACACC  NAAAACAAAA AAAAAAAAAA AAAAAAA
```

Seq ID NO: 300 Protein sequence
Protein Accession #: NP_003049.1

```
       1          11         21         31         41         51
       |          |          |          |          |          |
     MPTTVDDVLE HGGEFHFFQK QMFFLLALLS ATFAPIYVGI VFLGFTPDHR CRSPGVAELS    60
     LRCGWSPAEE LNYTVPGPGP AGEASPRQCR RYEVDWNQST FDCVDPLASL DTNRSRLPLG   120
     PCRDGWVYET PGSSIVTEFN LVCANSWMLD LFQSSVNVGF FIGSMSIGYI ADRFGRKLCL   180
65   LTTVLINAAA GVLMAISPTY TWMLIFRLIQ GLVSKAGWLI GYILITEFVG RRYRRTVGIF   240
     YQVAYTVGLL VLAGVAYALP HWRWLQPTVA LPNFFLLYY  WCIPESPRWL ISQNKNAEAM   300
     RIIKHIAKKN GKSLPASLQR LRLEEETGKK LNPSFIDLVR TPQIRKHTMI LMYNWFTSSV   360
     LYQGLIMHMG LAGDNIYLDF FYSALVEFPA AFMIILTIDR IGRRYPWAAS NMVAGAACLA   420
     SVFIPGDLQW LKIIISCLGR MGITMAYEIV CLVNAELYPT FIRNLGVHIC SSMCDIGGII   480
70   TPFLVYRLIN INLELPLMVF GVLGLVAGGL VLLLPETKGK ALPETIEEAE DMQRPRKNKE   540
     RMIYLQVQKL DIPLN
```

Seq ID NO: 301 DNA sequence
Nucleic Acid Accession #: NM_012206.1
Coding sequence: 52..1131

```
       1          11         21         31         41         51
       |          |          |          |          |          |
     GTTACCCAGC ATTGTGAGTG ACAGAGCCTG GATCTGAACG CTGATCCCAT AATGCATCCT    60
     CAAGTGGTCA TCTTAAGCCT CATCCTACAT CTGGCAGATT CTGTAGCTGG TTCTGTAAAG   120
80   GTTGGTGGAG AGGCAGGTCC ATCTGTCACA CTACCCTGACA ACTACAGTGG AGCTGTCACA   180
     TCAATGTGCT GGAATAGAGG CTCATGTTCT CTATTCACAT GCCAAAATGG CATTGTCTGG   240
     ACCAATGGAA CCCACGTCAC CTATCGGAAG GACACACGCT ATAAGCTATT GGGGGACCTT   300
     TCAAGAAGGG ATGTCTCTTT GACCATAGAA AATACAGCTG TGTCTGACAG TGGCGTATAT   360
```

```
TGTTGCCGTG TTGAGCACCG TGGGTGGTTC AATGACATGA AAATCACCGT ATCATTGGAG   420
ATTGTGCCAC CCAAGGTCAC GACTACTCCA ATTGTCACAA CTGTTCCAAC CGTCACGACT   480
GTTCGAACGA GCACCACTGT TCCAACGACA ACGACTGTTC CAACGACAAC TGTTCCAACA   540
ACAATGAGCA TTCCAACGAC AACGACTGTT CCGACGACAA TGACTGTTTC AACGACAACG   600
AGCGTTCCAA CGACAACGAG CATTCCAACA ACAACAAGTG TTCCAGTGAC AACAACGGTC   660
TCTACCTTTG TTCCTCCAAT GCCTTTGCCC AGGCAGAACC ATGAACCAGT AGCCACTTCA   720
CCATCTTCAC CTCAGCCAGC AGAAACCCAC CCTACGACAC TGCAGGGAGC AATAAGGAGA   780
GAACCCACCA GCTCACCATT GTACTCTTAC ACAACAGATG GGAATGACAC CGTGACAGAG   840
TCTTCAGATG GCCTTTGGAA TAACAATCAA ACTCAACTGT TCCTAGAACA TAGTCTACTG   900
ACGGCCAATA CCACTAAAGG AATCTATGCT GGAGTCTGTA TTTCTGTCTT GGTGCTTCTT   960
GCTCTTTTGG GTGTCATCAT TGCCAAAAAG TATTTCTTCA AAAAGGAGGT TCAACAACTA  1020
AGTGTTTCAT TTAGCAGCCT TCAAATTAAA GCTTTGCAAA ATGCAGTTGA AAAGGAAGTC  1080
CAAGCAGAAG ACAATATCTA CATTGAGAAT AGTCTTTATG CCACGGACTA AGACCCAGTG  1140
GTGCTCTTTG AGAGTTTACG CCCATGACTG CAGAAGACTG AACAGGTATC AGCACATCAG  1200
ATGTCTTTTA GACTCCAAGA CAATTTTTCT GTTTCAGTTT CATCTGGCAT TCCAACATGT  1260
CAGTGATACT GGGTAGAGTA ACTCTCCCAC TCCAAACTGT GTATAGTCAA CCTCATCATT  1320
AATGTAGTCC TAATTTGTTT TGCTAAAACT GGCTCAATCC TTCTGATCAT TGCAGAGTTT  1380
TCTCTCAAAC ATGAACACTT TAGAATTGTA TGTTCTCTTT AGACCCCATA AATCCTGTAT
```

Seq ID NO: 302 <u>Protein sequence</u>
Protein Accession #: NP_036338.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MHPQVVILSL ILHLADSVAG SVKVGGEAGP SVTLPCHYSG AVTSMCWNRG SCSLFTCQNG   60
IVWTNGTHVT YRKDTRYKLL GDLSRRDVSL TIENTAVSDS GVYCCRVEHR GWFNDMKITV  120
SLEIVPPKVT TTPIVTTVPT VTTVRTSTTV PTTTTVPTTT VPTTMSIPTT TTVPTTMTVS  180
TTTSVPTTTS IPTTTSVPVT TTVSTFVPPM PLPRQNHEPV ATSPSSPQPA ETHPTTLQGA  240
IRREPTSSPL YSYTTDGNDT VTESSDGLWN NNQTQLFLEH SLLTANTTKG IYAGVCISVL  300
VLLALLGVII AKKYFFKKEV QQLSVSFSSL QIKALQNAVE KEVQAEDNIY IENSLYATD
```

Seq ID NO: 303 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_001044.1
Coding sequence: 129..1991

```
1          11         21         31         41         51
|          |          |          |          |          |
ACCGCTCCGG AGCGGGAGGG GAGGCTTCGC GGAACGCTCT CGGCGCCAGG ACTCGCGTGC   60
AAAGCCCAGG CCCGGGCGGC CAGACCAAGA GGGAAGAAGC ACAGAATTCC TCAACTCCCA  120
GTGTGCCCAT GAGTAAGAGC AAATGCTCCG TGGGGACTCAT GTCTTCCGTG GTGGCCCCGG  180
CTAAGGAGCC CAATGCCGTG GGCCCGAAGG AGGTGGAGCT CATCCTTGTC AAGGAGCAGA  240
ACGGAGTGCA GCTCACCAGC TCCACCCTCA CCAACCCGCG GCAGAGCCCC GTGGGAGGCCC  300
AGGATCGGGA GACCTGGGGC AAGAAGATCG ACTTTCTCTC GTCCGTCATT GGCTTTGCTG  360
TGGACCTGGC CAACGTCTGG CGGTTCCCCT ACCTGTGCTA CAAAAATGGT GGCGGTGCCT  420
TCCTGGTCCC CTACCTGCTC TTCATGGTCA TTGCTGGGAT GCCACTTTTC TACATGGAGC  480
TGGCCCTGGG CCAGTTCAAC AGGGAAGGGG CCGCTGGTGT CTGGAAGATC TGCCCCATAG  540
TGAAAGGTGT GGGCTTCACG GTCATCCTCA TCTCACTGTA TGTCGGCTTC TTCTACAACG  600
TCATCATCGC CTGGGCGCTG CACTATCTCT TCTCCTCCTT CACCACGGAG CTCCCCTGGA  660
TCCACTGCAA CAACTCCTGG AACAGCCCGA ACTGCTCGGA TGCCCATCCT GGTGACCCAG  720
GTGGAGACAG CTCGGCCCTC AACGACACTT TTGGGACCAC ACCTGCTGCC GAGTACTTTG  780
AACGTGGCGT GCTGCACCTC CACCAGAGCC ATGGCATCGA CGACCTGGGG CCTCCGCGGT  840
GGCAGCTCAC AGCCTGCCTG GTGCTGGTCA TCGTGCTGCT CTACTTCAGC CTCTGGAAGG  900
GCGTGAAGAC CTCAGGGAAG GTGGTATGGA TCACAGCCAC CATGCCATAC GTGGTCCTCA  960
CTGCCCTGCT CCTGCGTGGG GTCACCCTCC TGGAGCCAT AGACGGCATC AGAGCATACC  1020
TGAGCGTTGA CTTCTACCGG CTCTGCGAGG CGTCTGTTTG GATTGACGCG GCCACCCAGG  1080
TGTGCTTCTC CCTGGGCGTG GGGTTCGGGG TGCTGATCGC CTTCTCCAGC TACAACAAGT  1140
TCACCAACAA CTGCTACAGG GACGCGGATTG TCACCACCTC CATCAACTCC CTGACGAGCT  1200
TCTCCTCCGG CTTCGTCGTC TTCTCCTTCC TGGGGTACAT GGCACAGAAG CACAGTGTGC  1260
CCATCGGGGA CGTGGCCAAG GACGGGCCAG GGCTGATCTT CATCATCTAC CCGGAAGCCA  1320
TCGCCACGCT CCCTCTGTCC TCAGCCTGGG CCGTGGTCTT CTTCATCATG CTGCTCACCC  1380
TGGGTATCGA CAGCGCCATG GGTGGTATGG AGTCAGTGAT CACCGGGCTC ATCGATGAGT  1440
TCCAGCTGCT GCACAGACAC CGTGAGCTCT TCACGCTCTT CATCGTCCTG GCGACCTTCG  1500
TCCTGTCCCT GTTCTGCGTC ACCAACGGTG GCATCTACGT CTTCACGCTC CTGGACCATT  1560
TTGCAGCCGG CACGTCCATC CTCTTTGGAG TGCTCATCGA AGCCATCGGA GTGGCCTGGT  1620
TCTATGGTGT TGGGCAGTTC ACGGACGACA TCCAGCAGAT GACCGGGCAG CGGCCCAGCC  1680
TGTACTGGCG GCTGTGCTGG AAGCTGGTCA GCCCCTGCTT TCTCCTGTTC GTGGTCGTGG  1740
TCAGCATTGT GACCTTCAGA CCCCCCCACT ACGGAGCCTA CATCTTCCCC GACTGGGCCA  1800
ACGCGCTGGG CTGGGTCATC GCCACATCCT CCATGGCCTT CATGTGGCCT CTGGGGGCCA  1860
ACAAGTTCTG CAGCCTGCCT GGGTCCTTTC GAGAGAAACT GGCCTACGCC ATTGCACCCG  1920
AGAAGGACCG TGAGCTGGTG GACAGAGGGG AGGTGCGCCA GTTCACGCTC CGCCACTGGC  1980
TCAAGGTGTA GAGGGAGCAG AGACGAAGAC CCCAGGAAGT CATCCTGCAA TGGGAGAGAC  2040
ACGAACAAAC CAAGGAAATC TAAGTTTCGA GAGAAAGGAG GGCAACTTCT ACTCTTCAAC  2100
CTCTACTGAA AACACAAACA ACAAAGCAGA AGACTCCTCT CTTCTGACTG TTTACACCTT  2160
TCCGTGCCGG GAGCGCACCT CGCCGTGTCT TGTGTTGCTG TAATAACGAC GTAGATCTGT  2220
GCAGCGGAGGT CCACCCCGTT GTTGTCCCTG CAGGGCAGAA AACGTCTAA CTTCATGCTG  2280
TCTGTGTGAG GCTCCCTCCC TCCCTGCTCC CTGCTCCCGG CTCTGAGGCT GCCCCAGGGG  2340
CACTGTGTTC TCAGGCGGGG ATCACGATCC TTGTAGACGC ACCTGCTGAG AATCCCCGTG  2400
CTCACAGTAG CTTCCTAGAC CATTTACTTT GCCCATATTA AAAAGCCAAG TGTCCTGCTT  2460
GGTTTAGCTG TGCAGAAGGT GAAATGGAGG AAACCACAAA TTCATGCAAA GTCCTTTCCC  2520
GATGCGTGGC TCCCAGCAGA GGCCGTAAAT TGAGCGTTCA GTTGACACAT TGCACACACA  2580
GTCTGTTCAG AGGCATTGGA GGATGGGGGT CCTGGTATGT CTCACCAGGA AATTCTGTTT  2640
ATGTTCTTGC AGCAGAGAGA AATAAAACTC CTTGAAACCA GCTCAGGCTA CTGCCACTCA  2700
GGCAGCCTGT GGGTCCTTGT GGTGTAGGGA ACGGCCTGAG AGGAGCGTGT CCTATCCCCG  2760
GACGCATGCA GGGCCCCCAC AGGAGCGTGT CCTATCCCCG GACGCATGCA GGGCCCCCAC  2820
```

```
AGGAGCATGT CCTATCCCTG GACGCATGCA GGGCCCCCAC AGGAGCGTGT ACTACCCCAG  2880
AACGCATGCA GGGCCCCCAC AGGAGCGTGT ACTACCCCAG GACGCATGCA GGGCCCCCAC  2940
TGGAGCGTGT ACTACCCCAG GACGCATGCA GGGCCCCCAC AGGAGCGTGT CCTATCCCCG  3000
GACCGGACGC ATGCAGGGCC CCCACAGGAG CGTGTACTAC CCCAGGACGC ATGCAGGGCC  3060
CCCACAGGAG CGTGTACTAC CCCAGGATGC ATGCAGGGCC CCCACAGGAG CGTGTACTAC  3120
CCCAGGACGC ATGCAGGGCC CCCATGCAGG CAGCCTGCAG ACCAACACTC TGCCTGGCCT  3180
TGAGCCGTGA CCTCCAGGAA GGGACCCCAC TGGAATTTTA TTTCTCTCAG GTGCGTGCCA  3240
CATCAATAAC AACAGTTTTT ATGTTTGCGA ATGGCTTTTT AAAATCATAT TTACCTGTGA  3300
ATCAAAACAA ATTCAAGAAT GCAGTATCCG CGAGCCTGCT TGCTGATATT GCAGTTTTTG  3360
TTTACAAGAA TAATTAGCAA TACTGAGTGA AGGATGTTGG CCAAAAGCTG CTTTCCATGG  3420
CACACTGCCC TCTGCCACTG ACAGGAAAGT GGATGCCATA GTTTGAATTC ATGCCTCAAG  3480
TCGGTGGGCC TGCCTACGTG CTGCCCGAGG GCAGGGGCCG TGCAGGGCCA GTCATGGCTG  3540
TCCCCTGCAA GTGGACGTGG GCTCCAGGGA CTGGAGTGTA ATGCTCGGTG GGAGCCGTCA  3600
GCCTGTGAAC TGCCAGGCAG CTGCAGTTAG CACAGAGGAT GGCTTCCCCA TTGCCTTCTG  3660
GGGAGGGACA CAGAGGACGG CTTCCCCATC GCCTTCTGGC CGCTGCAGTC AGCACAGAGA  3720
GCGGCTTCCC CATTGCCTTC TGGGGAGGGA CACAGAGGAC AGTTTCCCCA TCGCCTTCTG  3780
GTTGTTGAAG ACAGCACAGA GAGCGGCTTC CCCATCGCCT TCTGGGGAGG GGCTCCGTGT  3840
AGCAACCCAG GTGTTGTCCG TGTCTGTTGA CCAATCTCTA TTCAGCATCG TGTGGGTCCC  3900
TAAGCACAAT AAAAGACATC CACAATGAAA AAAAAAAAAG GAATTC
```

Seq ID NO: 304 Protein sequence
Protein Accession #: NP_001035.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MSKSKCSVGL MSSVVAPAKE PNAVGPKEVE LILVKEQNGV QLTSSTLTNP RQSPVEAQDR    60
ETWGKKIDFL LSVIGFAVDL ANVWRFPYLC YKNGGGAFLV PYLLFMVIAG MPLFYMELAL   120
GQFNREGAAG VWKICPILKG VGPTVILISL YVGFFYNVII AWALHYLFSS FTTELPWIHC   180
NNSWNSPNCS DAHPGDSSGD SSGLNDTFGT TPAAEYFERG VLHLHQSHGI DDLGPPRWQL   240
TACLVLVIVL LYFSLWKGVK TSGKVVWITA TMPYVVLTAL LLRGVTLPGA IDGIRAYLSV   300
DFYRLCEASV WIDAATQVCF SLGVGFGVLI AFSSYNKFTN NCYRDAIVTT SINSLTSFSS   360
GFVVFSFLGY MAQKHSVPIG DVAKDGPGLI FIIYPEAIAT LPLSSAWAVV FFIMLLTLGI   420
DSAMGGMESV ITGLIDEFQL LHRHRELFTL PIVLATPLLS LFCVTNGGIY VFTLLDHFAA   480
GTSILFGVLI EAIGVAWFYG VGQFSDDIQQ MTGQRPSLYW RLCWKLVSPC FLLFVVVVSI   540
VTFRPPHYGA YIFPDWANAL GWVIATSSMA MVPIYAAYKF CSLPGSFREK LAYAIAPEKD   600
RELVDRGEVR QFTLRHWLKV
```

Seq ID NO: 305 DNA sequence
Nucleic Acid Accession #: NM_001216.1
Coding sequence: 43..1422

```
1          11         21         31         41         51
|          |          |          |          |          |
GCCCGTACAC ACCGTGTGCT GGGACACCCC ACAGTCAGCC GCATGGCTCC CCTGTGCCCC    60
AGCCCCTGGC TCCCTCTGTT GATCCCGGCC CCTGCTCCAG GCCTCACTGT GCAACTGCTG   120
CTGTCACTGC TGCTTCTGAT GCCTGTCCAT CCCCAGAGGT TGCCCCGGAT GCAGGAGGAT   180
TCCCCCTTGG GAGGAGGCTC TTCTGGGGAA GATGACCCAC TGGGCGAGGA GGATCTGCCC   240
AGTGAAGAGG ATTCACCCAG AGAGGAGGAT CCACCCGGAG AGGAGGATCT ACCTGGAGAG   300
GAGGATCTAC CTGGAGAGGA GGATCTACCT GAAGTTAAGC CTAAAATCAGA AGAAGGAGGC   360
TCCCTGAAGT TAGAGGATCT ACCTACTGTT GAGGCTCCTG GAGATCCTCA AGAACCCCAG   420
AATAATGCCC ACAGGGACAA AGAAGGGGAT GACCAGAGTC ATTGGCGCTA TGGAGGCGAC   480
CCGCCCTGGC CCGGGGTGTC CCCAGCCTGC GCGGGCCGCT TCCAGTCCCC GGTGGATATC   540
CGCCCCCAGC TCGCCGCCTT CTGCCCGGCC CTGCGCCCCC TGGAACTCCT GGGCTTCCAG   600
CTCCCGCCGC TCCCAGAACT GCGCCTGCGC AACAATGGCC ACAGTGTGCA ACTGACCCTG   660
CCTCCTGGGC TAGAGATGGC TCTGGGTCCC GGGCGGGAGT ACCGGGCTCT GCAGCTGCAT   720
CTGCACTGGG GGGCTGCAGG TCGTCCGGGC TCGGAGCACA CTGTGGAAGG CCACCGTTTC   780
CCTGCCGAGA TCCACGTGGT TCACCTCAGC ACCGCCTTTG CCAGAGTTGA CGAGGCCTTG   840
GGGCGCCCGG GAGGCCTGGC CGTGTTGGCC GCCTTTCTGG AGGAGGGCCC GGAAGAAAAC   900
AGTGCCTATG AGCAGTTGCT GTCTCGCTTG GAAGAAATCG CTGAGGAAGG CTCAGAGACT   960
CAGGTCCCAG GACTGGACAT ATCTGCACTC CTGCCCTCTG ACTTCAGCCG CTACTTCCAA  1020
TATGAGGGGT CTCTGACTAC ACCGCCCTGT GCCCAGGGTG TCATCTGGAC TGTGTTTAAC  1080
CAGACAGTGA TGCTGAGTGC TAAGCAGCTC CACACCCTCT CTGACACCCT GTGGGGACCT  1140
GGTGACTCTC GGCTACAGCT GAACTTCCGA GCGACGCAGC CTTTGAATGG GCGAGTGATT  1200
GAGGCCTCCT TCCCTGCTGG AGTGGACAGC AGTCCTCGGG CTGCTGAGCC AGTCCAGCTG  1260
AATTCCTGCC TGGCTGCTGG TGACATCCTA GCCCTGGTTT TTGGCCTGCT TTTTGCTGTC  1320
ACCAGCGTCG CGTTCCTTGT GCAGATGAGA AGGCAGCACA GAAGGGGAAC CAAAGGGGGT  1380
GTGAGCTACC GCCCAGCAGA GGTAGCCGAG ACTGGAGCCT AGAGGCTGGA TCTTGGAGAA  1440
TGTGAGAAGC CAGCCAGAGG CATCTGAGGG GGAGCCGGTA ACTGTCCTGT CCTGCTCATT  1500
ATGCCACTTC CTTTTAACTG CCAAGAAATT TTTTAAAATA AATATTTATA AT
```

Seq ID NO: 306 Protein sequence
Protein Accession #: NP_001207.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MAPLCPSPWL FLLIPAPAPG LTVQLLLSLL LLMPVEPQRL PRMQEDSPLG GGSSGEDDPL    60
GEEDLPSEED SPREEDPPGE EDLPGEEDLP GEEDLPEVKP KSEEEGSLKL EDLPTVEAPG   120
DPQEFQNNAH RDKEGDDQSH WRYGGDPPWP RVSPACAGRF QSPVDIRPQL AAFCPALRPL   180
ELLGFQLPPL PELRLRNNGH SVQLTLPPGL EMALGPGREY RALQLHLHWG AAGRPGSEHT   240
VEGHRFPAEI HVVHLSTAFA RVDEALGRPG GLAVLAAFLE EGPEENSAYE QLLSRLEEIA   300
EEGSETQVPG LDISALLPSD FSRYFQYEGS LTTPPCAQGV INTVFNQTVM LSAKQLHTLS   360
DTLWGPGDSR LQLNFRATQP LNGRVIEASF PAGVDSSPRA AEPVQLNSCL AAGDILALVF   420
GLLFAVTSVA FLVQMRRQHR RGTKGGVSYR PAEVAETGA
```

Seq ID NO: 307 DNA sequence
Nucleic Acid Accession #: NM_003039.1
Coding sequence: 76..1581

```
1          11         21         31         41         51
|          |          |          |          |          |
CTTCTCTCTC CATTCAGTGC ACGCGTTACT TTGGCTAAAA GGAGGTGAGC GGCACTCTGC   60
CCTTCCAGAG CAAGCATGGA GCAACAGGAT CAGAGCATGA AGGAAGGGAG GCTGACGCTT  120
GTGCTTGCCC TGGCAACCCT GATAGCTGCC TTTGGGTCAT CCTTCCAGTA TGGGTACAAC  180
GTGGCTGCTG TCAACTCCCC AGCACTGCTC ATGCAACAAT TTTACAATGA GACTTACTAT  240
GGTAGGACCG GTGAATTCAT GGAAGACTTC CCCTTGACGT TGCTGTGGTC TGTAACCGTG  300
TCCATGTTTC CATTTGGAGG GTTTATCGGA TCCCTCCTGG TCGGCCCCTT GGTGAATAAA  360
TTTGGCAGAA AAGGGGCCTT GCTGTTCAAC AACATATTTT CTATCGTGCC TGCGATCTTA  420
ATGGGATGCA GCAGAGTCGC CACATCATTT GAGCTTATCA TTATTTCCAG ACTTTTGGTG  480
GCAATATGTG CAGGTGTATC TTCCAACGTG GTCCCCATGT ACTTAGGGGA GCTGGCCCCT  540
AAAAACCTGC GGGGGGCTCT CGGGGTGGTG CCCCAGCTCT TCATCACTGT TGGCATCCTT  600
GTGGCCCAGA TCTTTGGTCT TCGGAATCTC CTTGCAAACG TAGATGGCTG GCCGATCCTG  660
CTGGGGCTGA CCGGGATCCC CGCGGCGCTG CAGCTCCTTC TGCTGCCCTT CTTCCCCGAG  720
AGCCCCAGGT ACCTGCTGAT TCAGAAGAAA GACGAAGCGG CCGCCAAGAA AGCCCTACAG  780
ACGCTGCGCG GCTGGGACTC TGTGGACAGG GAGGTGGCCG AGATCCGGCA GGAGGATGAG  840
GCAGAGAAGG CCGCGGGCTT CATCTCCGTG CTGAAGCTGT TCCGGATGCG CTCGCTGCGC  900
TGGCAGCTGC TGTCCATCAT CGTCCTCATG GGCGGCCAGC AGCTGTCGGG CGTCAACGCT  960
ATCTACTACT ACGCGGACCA GATCTACCTG AGCGCCGGCG TGCCGGAGGA GCACGTGCAG 1020
TACGTGACGG CCGGCACCGG GGCCGTGAAC GTGGTCATGA CCTTCTGCGC CGTGTTCGTG 1080
GTGGAGCTCC TGGGTCGGAG GCTGCTGCTG CTGCTGGGCT TCTCCATCTG CCTCATAGCC 1140
TGCTGCGTGC TCACTGCAGC TCTGGCACTG CAGGACACAG TGTCCTGGAT GCCATACATC 1200
AGCATCGTCT GTGTCATCTC CTACGTCATA GGACATGCCG TCGGGCCCAG TCCCATACCC 1260
GCGCTGCTCA TCACTGAGAT CTTCCTGCAG TCCTCTCGGC CATCTGCCTT CATGGTGGGG 1320
GGCAGTGTGC ACTGGCTCTC CAACTTCACC GTGGGCTTGA TCTTCCCGTT CATCCAGGAG 1380
GGCCTCGGCC CGTACAGCTT CATTGTCTTC GCCGTGATCT GCCTCCTCAC CACCATCTAC 1440
ATCTTCTTGA TTGTCCCGGA GACCAAGGCC AAGACGTTCA TAGAGATCAA CCAGATTTTC 1500
ACCAAGATGA ATAAGGTGTC TGAAGTGTAC CCGGAAAAGG AGGAACTGAA AGAGCTTCCA 1560
CCTGTCACTT CGGAACAGTG ACTCTGGAGA GGAAGCCAGT GGAGCTGGTC TGCCAGGGGC 1620
TTCCCACTTT GGCTTATTTT TCTGACTTCT AGCTGTCTGT GAATATCCAG AAATAAAACA 1680
ACTCTGATGT GGAATGCAGT CCTCATCTCC AGCCTCCCCA CCCCAGTGGG AACTGTGCAA 1740
AGGGCTGCCT TGCTGTTCTT GAAGCTGGGC TGTCTCTCTC CATGTTGGCC TGTCACCAGA 1800
CCCGAGTCAA TTAAACAGCT GGTCCTCCAC TTTGCTGGTT CAGCCTTCGT GTGGCTCCTG 1860
GTAACGTGGC TCCACCTGA TGGGTCAACC TTTGTGTGGC TCCTGGTAAC ATAACAACAA 1920
CAGTTACTAT AGTGGTGAGA TGGAAGGAAT CAAATTTTGC CAGAGAAACT AACTCGGTGG 1980
CCCCAACAGG TCTTCCGGGG CCATGGGCAT TTGTTTAGAG CCAAATTCAT CCTCTTACCA 2040
GATCCTTTTC CAGAAATACC TGTCTAGGAA GGTGTGATGT CAGAAACAAT GACATCCAGA 2100
AAGCTGAGGA ACAGGTTCCT GTGGAGACAC TGAGTCAGAA TTCTTCATCC AAATTATTTT 2160
GTTAGTGGAA AATGGAATTG CTTCTGTGTA GTCAATAAAA TGAACCTGAT CACTTTTC
```

Seq ID NO: 308 Protein sequence
Protein Accession #: NP_003030.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MEQQDQSMKE GRLTLVLALA TLIAAFGSSF QYGYNVAAVN SPALLMQQFY NETYYGRTGE   60
FMEDFPLTLL WSVTVSMFPP GGFIGSLLVG PLVNKFGRKG ALLFNNIFSI VPAILMGCSR  120
VATSFELIII SRLLVGICAG VSSNVVPMYL GELAPKNLRG ALGVVPQLFI TVGILVAQIF  180
GLRNLLANVD GWPILLGLTG VPAALQLLLL PFFPESPRYL LIQKEDEAAA KKALQTLRGW  240
DSVDREVAEI RQEDEAEKAA GFISVLKLFR MRSLRWQLLS IIVLMGGQQL SGVNAIYYYA  300
DQIYLSAGVP EEHVQYVTAG TGAVNVVMTF CAVFVVELLG RRLLLLLGFS ICLIACCVLT  360
AALALQDTVS WMPYISIVCV ISYVIGHALG PSPIPALLIT EIFLQSSRPS AFMVGGSVHW  420
LSNFTVGLIF PFIQEGLGPY SFIVFAVICL LTTIYIFLIV PETKAKTPIE INQIFTKMNK  480
VSEVYPEKEE LKELPPVTSE Q
```

Seq ID NO: 309 DNA sequence
Nucleic Acid Accession #: NM_001252.1
Coding sequence: 138..719

```
1          11         21         31         41         51
|          |          |          |          |          |
GGCTGGTCCC CTGACAGGTT GAAGCAAGTA GACGCCCAGG AGCCCGGGGA GGGGGCTGCA   60
GTTTCCTTCC TTCCTTCTCG GCAGCGCTCC GCGCCCCCAT CGCCCCTCCT GCGCTAGCGG  120
AGGTGATCGC CGCGGCGATG CCGGAGGGAGG GTTCGGGCTG CTCGGTGCGG CGCAGGCCGT  180
ATGGGTGCGT CCTGCGGGCT GCTTTGGTCC CATTGGTCGC GGGCTTGGTG ATCGGCCTCG  240
TGGTGTGCAT CCAGCGCTTC GCACAGGCTC AGCAGCAGCT GCCGCTCGAG TCACTTGGGT  300
GGGACGTAGC TGAGCTGCAG CTGAATCACA CAGGACCTCA GCAGGACCCC AGGCTATACT  360
GGCAGGGGGG CCCAGCACTG GGCCGCTCCT TCCTGCATGG ACCAGAGCTG GACAAGGGGC  420
AGCTACGTAT CCATCGTGAT GGCATCTACA TGGTACACAT CCAGGTGACG CTGGCCATCT  480
GCTCCTCCAC GACGGCCTCC AGGCACCACC CCACCACCCT GGCCGTGGGA ATCTGCTCTC  540
CCGCCTCCCG TAGCATCAGC CTGCTGCGTC TCAGCTTCCA CCAAGGTTGT ACCATTGCCT  600
CCCAGCCGCCT GACGCCCCTG GCCCGAGGGG ACACACTCTG CACCAACCTC ACTGGGACAC  660
TTTTGCCTTC CCGAAACACT GATGAGACCT TCTTTGGAGT GCAGTGGGTG CGCCCCTGAC  720
CACTCGTGCT GATTAGGGTT TTTTAAATTT TATTTTATTT TATTTAAGTT CAAGAGAAAA  780
AGTGTACACA CAGGGGCCAC CCGGGGTTGG GGTGGGAGTG TGGTGGGGGG TAGTGGTGGC  840
AGGACAAGAG AAGGCATTGA GCTTTTTCTT TCATTTTCCT ATTAAAAA
```

Seq ID NO: 310 Protein sequence
Protein Accession #: NP_001243.1

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       MPEEGSGCSV RRRPYGCVLR AALVPLVAGL VICLVVCIQR FAQAQQQLPL ESLGWDVAEL   60
       QLNHTGPQQD PRLYWQGGPA LGRSFLHGPE LDKGQLRIHR DGIYMVHIQV TLAICSSTTA   120
       SRHHPTTLAV GICSPASRSI SLLRLSFEQG CTIASQRLTP LARGDTLCTN LTGTLLPSRN   180
       TDETFPGVQW VRP
```

Seq ID NO: 311 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..3978

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       ATGGTGGGTG AAGGACCCTA CCTTATCTCA GATCTGGACC AGCGAGGCCG GCGGAGATCC   60
       TTTGCAGAAA GATATGACCC CAGCCTGAAG ACCATGATCC CAGTGCGACC CTGTGCAAGG   120
       TTAGCACCCA ACCCGGTGGA TGATGCCGGG CTACTCTCCT TCGCCACATT TTCCTGGCTC   180
       ACGCCGGTGA TGGTGAAAGG CTACCGGCAA AGGCTGACCG TAGACACCCT GCCCCCCATTG   240
       TCGACATATG ACTCATCTGA CACCAATGCC AAAAGATTTC GAGTCCTTTG GGATGAAGAG   300
       GTAGCAAGGG TGGGTCCTGA GAAGGCCTCT CTGAGCCACG TGGTGTGGAA ATTCCAGAGG   360
       ACACGCGTGT TGATGGACAT CGTGGCCAAC ATCCTGTGCA TCATCATGGC AGCCATAGGG   420
       CCGACAGTTC TCATTCACCA AATCCTCCAG CAGACTGAGA GGACCTCTGG GAAAGTCTGG   480
       GTTGGCATTG GACTGTGCAT AGCCCTTTTT GCCACCGAGT TTACCAAAGT CTTCTTTTGG   540
       GCCCTTGCCT GGGCCATCAA CTACCGCACG GCCATCCGGT TGAAGGTGGC GCTCTCCACC   600
       TTGGTTTTTG AAAACCTAGT GTCCTTCAAG ACATTGACCG ACATCTCTGT TGGCGAGGTG   660
       CTCAATATAC TGTCAAGTGA TAGCTATTCT TTGTTTGAAG CTGCCTTGTT TTGTCCTTTG   720
       CCAGCCACCA TCCCGATCCT AATGGTCTTT TGTGCGGCGT ACGCCTTTTT CATTCTGGGG   780
       CCCACAGCTC TCATCGGGAT ATCAGTGTAT GTCATATTCA TACCCGTCCA GATGTTTATG   840
       GCCAAGCTCA ATTCAGCTTT CCGAAGGTCA GCAATTTTGG TGACAGACAA GCGAGTTCAG   900
       ACAATGAATG AGTTTCTGAC CTGCATCAGG CTGATCAAAA TGTATGCCTG GGAGAAATCT   960
       TTTACCAACA CTATCCAAGA TATAAGAAGG AGGGAAAGAA AATTACTGGA AAAAGCTGGA   1020
       TTTGTCCAAA GTGGAAACTC TGCCCTGGCC CCCATCGTGT CCACCATAGC CATCGTGCTG   1080
       ACATTATCCT GCCACATCCT CCTGAGACGC AAACTCACCG CACCCGTGGC ATTTAGTGTG   1140
       ATTGCCATGT TTAATGTAAT GAAGTTTTCC ATTGCAATCT TGCCCTTCTC CATCAAAGCA   1200
       ATGGCTGAAG CGAATGTCTC TCTAAGGAGA ATGAAGAAAA TTCTCATAGA TAAAAGCCCC   1260
       CCATCTTACA TCACCCAACC AGAAGACCCA GATACTGTCT TGCTTTTAGC AAATGCCACC   1320
       TTGACATGGG AGCATGAAGC CAGCAGGAAA AGTACCCCAA AGAAATTGCA GAACCAGAAA   1380
       AGGCATTTAT GCAAGAAACA GAGGTCAGAG GCATACAGTG AGAGGAGTCC ACCAGCCAAG   1440
       GGAGCCACTG GCCCAGAGGA GCAAAGTGAC AGCCTCAAAT CGGTTCTGCA CAGCATAAGC   1500
       TTTGTGGTGA GAAAGTTATG TCGTTATCCC GAAGCCCAGC TCCTGGCTTG GAGGTGGCCA   1560
       GCAGTGTTTG TTGGGAGAAT CATCAGAGGA TACAGGCCTC ATGGATTTTC TGCTAAAGAC   1620
       AAGGATGAAT CTAGAAGGCT TCTTACTTGG CCCCAAGAAG TGGATAGGAC TCAAAGGGCA   1680
       GCCAAATACC TGGGGAAGAT CTTGGGAATA TGTGGGAATG TGGGAAGTGG AAAGAGCTCC   1740
       CTCCTTGCAG CTCTCCTAGG ACAGATGCAG CTGCAGAAAG GGGTGGTGCC AGTCAATGGA   1800
       ACTTTGGCCT ACGTTTCACA GCAGGCATGG ATCTTTCATG GAAATGTGAG AGAAAACATA   1860
       CTCTTTGGAA AAAAGTATGA TCACCAAAGG TATCAGCACA CAGTCCGGCGT CTGTGGCCTC   1920
       CAGAAGGACC TGAGCAACCT CCCCTATGGA GACCTGACTG AGATTGGGGA GCGGGGGCCTC   1980
       AACCTCTCTG GGGGGCAGAG GCAGAGGATT AGCCTGGCCC GCGCTGTCTA CTCCGACCGT   2040
       CAGCTCTACC TGCTGGACGA CCCCCTGTCG GCCGTGGACG CCCACGTGGG GAAGCACGTC   2100
       TTTGAGGAGT GCATTAAGAA CACGCTCAGG GGAAAGACAG TCGTCCTGGT GACCCACCAG   2160
       CTACAGTTCT TAGAGTCCTTG TGATGAAGTT ATTTTATTAG AAGATGGAGA GATTTGTGAA   2220
       AAGGGAACCC ACAAGGAGTT AATGGAGGAG AGAGGGCGCT ATGCAAAACT GATTCACAAC   2280
       CTGCGAGGAT TGCAGTTCAA GGATCCTGAA CACCTTTACA ATGCAGCAAT GGTGGAAGCC   2340
       TTCAAGGAGA GCCCTGCTGA GAGAGAGGAA GATGCTGGTA TAATCGGGTA CCTCCTTTCT   2400
       CTCTTCACTG TGTTCCTCTT CCTCCTGATG ATTGGCAGCG CTGCCCTTCAG CAACTGGTGG   2460
       CTGGGTCTCT GGTTGGACAA GGGCTCACGG ATGACCTGGT GGCCCCAGGG CAACAGGACC   2520
       ATGTGTGAGG TCGGCGCGGT GCTGGCAGAC ATCGGTCAGC ATGTGTACCA GTGGGTGTAC   2580
       ACTGCAAGCA TGGTGTTCAT GCTGGTGTTT GGCGTCACCA AAGGCTTCGT CTTCACCAAG   2640
       ACCACACTGA TGGCATCCTC CTCTCTGCAT GACACGGTGT TTGATAAGAT CTTAAAGAGC   2700
       CCAATGAGTT TCTTTGACAC GACTCCCACT GGCAGGCTAA TGAACCGGTTT TTCCAAGGAT   2760
       ATGGACGAGC TGGATGTGAG GCTGCCCGTTT CACGCAGAGA ACTTTCTGCA GCAGTTTTTT   2820
       ATGTGGGTGT TTATTCTCGT GATCTTGGCT GCTGTGTTTC CTGCTGTCCT TTTAGTCGTG   2880
       GCCAGCCTTG CTGTAGGCTT CTTCATTCTG TTACGCATTT CCACAGAGG AGTCCAGGAG   2940
       CTCAAGAAGG TGGAGAATGT CAGCCGGTCA CCCTGGTTCA CCCACATCAC CTCCTCCATG   3000
       CAGGGCCTGG GCATCATTCA CGCCTATGGC AAGAAGGAGA GCTGCATCAC CTATACTTCA   3060
       TCCAAAGGCC TGTCATTGTC ATACATCATC CAGCTGAGCG GACTGCTTCA AGTGTGTGTG   3120
       CGAACGGGAA CAGAGACGCA AGCCAAATTC ACCTCCGTGG AGCTGCTCAG GGAATACATT   3180
       TCGACCTGTG TTCCTGAATG CACTCATCCC CTCAAAGTGG GGACCTGTCC CAAGGACTGG   3240
       CCCAGCTGTG GGGAGATCAC CTTCAGAGAC TATCAGATGA GATACAGAGA CAACACCCCG   3300
       CTTGTTCTCG ACAGCCTGAA CTTGAACATA CAAAGTGGGC AGACAGTCGG GATTGTTGGA   3360
       AGAACAGGTT CCGGAAAGTC ATCGTTAGGA ATGGCTTTGT TTCGTCTGGT GGAGCCAGCC   3420
       AGTGGCACAA TCTTTATTGA TGAGGTGGAT ATCTGCATTC TCAGCTTGGA AGACCCAGA   3480
       ACCAAGCTGA CTGTGATCCC ACAGGATCCT GTCCTGTTTG TAGGTACAGT AAGGTACAAC   3540
       TTGGATCCCT TTGAGAGTCA CACCGATGAG ATGCTCTGGC AGGTTCTGGA GAGAACATTC   3600
       ATGAGAGACA CAATAATGAA ACTCCCAGAA AAATTACAGG CAGAAGTCAC AGAAAATGGA   3660
       GAAAACTTCT CAGTAGGGGA ACGTCAGCTG CTTTGTGTGG CCCGAGCTCT TCTCCGTAAT   3720
       TCAAGGATCA TTCTCCTTGA TGAAGCCACC GCCTCTATGG ACTCCAAGAC TGACACCCTG   3780
       GTTCAGAACA CCATCAAAGA TGCCTTCAAG GGCTGCACTG TGCTGACCAT CGCCCACCGC   3840
       CTCAACACAG TTCTCAACTG CGATCACGTC CTGGTTATGG AAAATGGGAA GGTGATTGAG   3900
       TTTGACAAGC CTGAAGTCCT TGCAGAGAAG CCAGATTCTG CATTTGCGAT GTTACTAGCA   3960
       GCAGAAGTCA GATTGTAG
```

Seq ID NO: 312 Protein sequence
Protein Accession #: Eos sequence

```
      1         11        21        31        41        51
      |         |         |         |         |         |
      MVGEGPYLIS DLDQRGRRRS FAERYDPSLK TMIPVRPCAR LAPNPVDDAG LLSFATFSWL   60
      TPVMVKGYRQ RLTVDTLPPL STYDSSDTNA KRFRVLWDEE VARVGPEKAS LSHVVWKFQR  120
      TRVLMDIVAN ILCIIMAAIG PTVLIHQILQ QTERTSGKVW VGIGLCIALF ATEFTKVFFW  180
      ALAWAINYRT AIRLKVALST LVFENLVSFX TLTHISVGEV LNILSSDSYS LPEAALFCPL  240
      PATIPILMVF CAAYAPPILG PTALIGISVY VIFIPVQMFM AKLNSAFRRS AILVTDKRVQ  300
      TMNEFLTCIR LIKMYAWEKS FTNTIQDIRR RERKLLEKAG FVQSGNSALA PIVSTIAIVL  360
      TLSCHILLRR KLTAPVAPSV IAMFNVMKFS IAILPFSIKA MAEANVSLRR MKKILIDKSP  420
      PSYITQPEDP DTVLLLANAT LTWEHEASRK STPKKLQNQK RHLCKKQRSE AYSERSPPAK  480
      GATGPEEQSD SLKSVLHSIS FVVRKLCRYP EAQLLAWRWP AVFVGRIIRG YRPHGFSAKD  540
      KDESRRLLTW PQEVDRTQRA AKYLGKILGI CGNVGSGKSS LLAALLGQMQ LQKGVVAVNG  600
      TLAYVSQQAW IFHGNVRENI LFGEKYDHQR YQHTVRVCGL QKDLSNLPYG DLTEIGERGL  660
      NLSGGQRQRI SLARAVYSDR QLYLLDDPLS AVDAEVGKEV FEECIKKTLR GKTVVLVTBQ  720
      LQFLESCDEV ILLEDGEICE KGTHKELMEE RGRYAKLIEN LRGLQFKDPE HLYNAAMVEA  780
      FKESPAEREE DAGIIGYLLS LFTVPLFLLM IGSAAFSNWW LGLWLDKGSR MTCGPQGNRT  840
      MCEVGAVLAD IGQHVYQWVY TASMVFMLVF GVTKGFVFTK TTLMASSSLH DTVFDKILKS  900
      PMSFFDTTPT GRLMNRFSKD MDELDVRLPF HAENFLQQFF MVVFILVILA AVFPAVLLVV  960
      ASLAVGPFIL LRIFHRGVQE LKKVENVSRS PWPTHITSSM QGLGIIHAYG KKESCITYTS 1020
      SKGLSLSYII QLSGLLQVCV RTGTETQAKP TSVELLREYI STCVPECTHP LKVGTCPKDW 1080
      PSCGEITFRD YQMRYRDNTP LVLDSLNLNI QSGQTVGIVG RTGSGKSSLG MALFRLVEPA 1140
      SGTIFIDEVD ICILSLEDLR TKLTVIPQDP VLFVGTVRYN LDPFESHTDE MLWQVLERTF 1200
      MRDTIMKLPE KLQAEVTENQ ENFSVGERQL LCVARALLRN SKIILLDEAT ASMDSKTDTL 1260
      VQNTIKDAFK GCTVLTIAHR LNTVLNCDHV LVMENGKVIE FDRPEVLAEK PDSAFAMLLA 1320
      AEVRL
```

Seq ID NO: 313 <u>DNA sequence</u>
Nucleic Acid Accession #: Z31560
Coding sequence: 1-966

```
      1         11        21        31        41        51
      |         |         |         |         |         |
      CACAGCGCCC GCATGTACAA CATGATGGAG ACGGAGCTGA AGCCGCCGGG CCCGCAGCAA   60
      ACTTCGGGGG GCGGCGGCGG CAACTCCACC GCGGCGGCGG CCGGCGGCAA CCAGAAAAAC  120
      AGCCCGGACC GCGTTCAAGCG GCCCATGAAT GCCTTCATGG TGTGGTCCCG CGGGCAGCGG  180
      CGCAAGATGG CCCAGGAGAA CCCCAAGATG CACAACTCGG AGATCAGCAA GCGCGTGGGC  240
      GCCGAGTGGA AACTTTTGTC GGAGACGGAG AAGCGGCCGT TCATCGACGA GGCTAAGCGG  300
      CTGCGAGCGC TGCACATGAA GGAGCACCCG GATTATAAAT ACCGGCCCCG GCGGAAAACC  360
      AAGACGCTCA TGAAGAAGGA TAAGTACACG CTGCCCGGCG GGCTGCTGGC CCCCGGCGGC  420
      AATAGCATGG CGAGCGGGGT CGGGGTGGGC GCCGGCCTGG GCGGGGCGGT GAACCAGCGC  480
      ATGGACAGTT ACGCGCACAT GAACGGCTGG AGCAACGGCA GCTACAGCAT GATGCAGGAC  540
      CAGCTGGGCT ACCCGCAGCA CCCGGGCCTC AATGCGCACG GCGCAGCGCA GATGCAGCCC  600
      ATGCACCGCT ACGACGTGAG CGCCCTGCAG TACAACTCCA TGACCAGCTC GCAGACCTAC  660
      ATGAACGGCT CGCCCACCTA CAGCATGTCC TACTCGCAGC AGGGCACCCC TGGCATGGCT  720
      CTTGGCTCCA TGGGTTCGGT GGTCAAGTCC GAGGCCAGCT CCAGCCCCCC TGTGGTTACC  780
      TCTTCCTCCC ACTCCAGGCC GCCCTGCCAG GCCGGGGACC TCCGGGACAT GATCAGCATG  840
      TATCTCCCCG GCGCCGAGGT GCCGGAACCC GCCGCCCCCA GCAGACTTCA CATGTCCCAG  900
      CACTACCAGA GCGGGCCCGGT GCCCGGCACG GCCATTAACG GCACACTGCC CCTCTCACAC  960
      ATGTGAGGGC CGGACAGCGA ACTGGAGGGG GGAGAAATTT CAAAGAAAA ACGAGGGAAA 1020
      TGGGAGGGGT GCAAAGAGG AGAGTAAGAA ACAGCATGGA GAAACCCGG TACGCTCAAA 1080
      AAAAA
```

Seq ID NO: 314 <u>Protein sequence</u>
Protein Accession #: CAA83435

```
      1         11        21        31        41        51
      |         |         |         |         |         |
      HSARMYNMME TELKPPGPQQ TSGGGGGNST AAAAGGNQKN SPDRVKRPMN AFMVWSRGQR   60
      RKMAQENPKM HNSEISKRLG AEWKLLSETE KRPPIDEAKR LRALEMKEHP DYKYRPRRKT  120
      KTLMKKDKYT LPGGLLAPGG NSMASGVGVG AGLGAGVNQR MDSYAHMNGW SNGSYSMMQD  180
      QLGYPQHPGL NAHGAAQMQP MHRYDVSALQ YNSMTSSQTY MNGSPTYSMS YSQQGTPGMA  240
      LGSMGSVVKS EASSSPPVVT SSSHSRAPCQ AGDLRDMISM YLPGAEVPEP AAPSRLHMSQ  300
      HYQSGPVPGT AINGTLPLSH M
```

Seq ID NO: 315 <u>DNA sequence</u>
Nucleic Acid Accession #: U91618
Coding sequence: 29..541

```
      1         11        21        31        41        51
      |         |         |         |         |         |
      CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAAATCC AGCTTGTATG   60
      CATGCTACTC CTGGCTTTCA GCTCCTGGAG TCTGTGCTCA GATTCAGAAG AGGAAATGAA  120
      AGCATTAGAA GCAGATTTCT TGACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT  180
      TCCCTCTTGG AAGATGACTC TGCTAAATGT TTGCAGTCTT GTAAATAATT TGAACAGCCC  240
      AGCTGAGGAA ACAGGAGAAG TTCATGAAGA GGAGCTTGTT GCAAGAAGGA AACTTCCTAC  300
      TGCTTTAGAT GGCTTTAGCT TGGAAGCAAT GTTGACAATA TACCAGCTCC ACAAAATCTG  360
      TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCTTG ATACTGGAAA  420
      TGACAAAAAT GGAAAGGAAG AAGTCATAAA GAGAAAAATT CCTTATATTC TGAAACGGCA  480
      GCTGTATGAG AATAAACCCA GAAGACCCTA CATACTCAAA AGAGATTCTT ACTATTACTG  540
      AGAGAATAAA TCATTTATTT ACATGTGATT GTGATTCATC ATCCCTTAAT TAAATATCAA  600
      ATTATATTTG TGTGAAAATG TGACAAACAC ACTTATCTGT CTCTTCTACA ATTGTGGTTT  660
      ATTGAATGTG TTTTTCTGCA CTAATAGAAA TTAGACTAAG TGTTTTCAAA TAAATCTAAA  720
      TCTTCAAAAA AAAAAAAAA AAATGGGGCC GCAATT
```

Seq ID NO: 316 Protein sequence
Protein Accession #: AAB50564

```
     1         11        21        31        41        51
     |         |         |         |         |         |
     MMAGMKIQLV CMLLLAFSSW SLCSDSEEEM KALEADFLTN MHTSKISKAH VPSWKMTLLN    60
     VCSLVNNLNS PAEETGEVHE EELVARRKLP TALDGFSLEA MLTIYQLRKI CHSRAFQHWE   120
     LIQEDILDTG NDKNGKEEVI KRKIPYILKR QLYENKPRRP YILKRDSYYY
```

Seq ID NO: 317 DNA sequence
Nucleic Acid Accession #: NM_006536.2
Coding sequence: 109..2940

```
     1         11        21        31        41        51
     |         |         |         |         |         |
     ACCTAAAACC TTGCAAGTTC AGGAAGAAAC CATCTGCATC CATATTGAAA ACCTGACACA    60
     ATGTATGCAG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACAACAT GACCCAAAGG   120
     AGCATTGCAG GTCCTATTTG CAACCTGAAG TTTGTGACTC TCCTCGGTTGC CTTAAGTTCA   180
     GAACTCCCAT TCCTGGGAGC TGGAGTACAG CTTCAAGACA ATGGGTATAA TGGATTGCTC   240
     ATTGCAATTA ATCCTCAGGT ACCTGAGAAT CAGAACCTCA TCTCAAACAT TAAGGAAATG   300
     ATAACTGAAG CTTCATTTTA CCTATTTAAT GCTACCAAGA GAAGAGTATT TTTCAGAAAT   360
     ATAAAGATTT TAATACCTGC CACATGGAAA GCTAATAATA ACAGCAAAAT AAAACAAGAA   420
     TCATATGAAA AGGCAAATGT CATAGTGACT GACTGGTATG GGGCACATGG AGATGATCCA   480
     TACACCCTAC AATACAGAGG GTGTGGAAAA GAGGGAAAT ACATTCATTT CACACCCTAAT   540
     TTCCTACTGA ATGATAACTT AACAGCTGGC TACGGATCAC GAGGCCGAGT GTTTGTCCAT   600
     GAATGGGCCC ACCTCCGTTG GGGTGTGTTC GATGAGTATA ACAATGACAA ACCTTTCTAC   660
     ATAAATGGGC AAAATCAAAT TAAAGTGACA AGGTGTTCAT CTGACATCAC AGGCATTTTT   720
     GTGTGTGAAA AAGGTCCTTG CCCCCAAGAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA   780
     GGATGCACCT TTATCTACAA TAGCACCCAA AATGCAACTG CATCAATAAT GTTCATGCAA   840
     AGTTTATCTT CTGTGGTTGA ATTTTGTAAT GCAAGTACCC ACAACCAAGA AGCACCAAAC   900
     CTACAGAACC AGATGTGCAG CCTCAGAAGT GCATGGGATG TAATCACAGA CTCTGCTGAC   960
     TTTCACCACA GCTTTCCCAT GAATGGGACT GAGCTTCCAC CTCCTCCCAC ATTCTCGCTT  1020
     GTACAGGCTG GTGACAAAGT GGTCTGTTTA GTGCTGGATG TGTCCAGCAA GATGGCAGAG  1080
     GCTGACAGAC TCCTTCAACT ACAACAAGCC GCAGAATTTT ATTTGATGCA GATTGTTGAA  1140
     ATTCATACCT TCGTGGGCAT TGCCAGTTTC GACAGCAAAG GAGAGATCAG AGCCCAGCTA  1200
     CACCAAATTA ACAGCAATGA TGATCGAAAG TTGCTGGTTT CATATCTGCC CACCACTGTA  1260
     TCAGCTAAAA CAGACATCAG CATTTGTTCA GGGCTTAAGA AAGGATTTGA GGTGGTTGAA  1320
     AAACTGAATG GAAAAGCTTA TGGCTCTGTG ATGATATTAG TGACCAGCGG AGATGATAAG  1380
     CTTCTTGGCA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAATTCA CTCCATTGCC  1440
     CTGGGTTCAT CTGCAGCCCC AAATCTGGAG GAATTATCAC GTCTTACAGG AGGTTTAAAG  1500
     TTCTTTGTTC CAGATATATC AAACTCCAAT AGCATGATTG ATGCTTTCAG TAGAATTTCC  1560
     TCTGGAACTG GAGACATTTT CCAGCAACAT ATTCAGCTTG AAAGTACAGG TGAAAATGTC  1620
     AAACCTCACC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT  1680
     ATGTTTCTAG TTACGTGGCA GGCCAGTGGT CCTCCTGAGA TTATATTATT TGATCCTGAT  1740
     GGACGAAAAT ACTACACAAA TAATTTTATC ACCAATCTAA CTTTTCGGAC AGCTAGTCTT  1800
     TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCCTGAACAA TACCCATCAT  1860
     TCTCTGCAAG CCCTGAAAGT GACAGTGACC TCTCGCGCCT CCAACTCAGC TGTGCCCCCA  1920
     GCCACTGTGG AAGCCTTTGT GGAAAGAGAC AGCCTCCATT TTCCTCATCC TGTGATGATT  1980
     TATGCCAATG TGAAACAGGG ATTTTATCCC ATTCTTAATG CCACTGTCAC TGCCACAGTT  2040
     GAGCCAGAGA CTGGAGATCC TGTTACGCTG AGACTCCTTG ATGATGGAGC AGGTGCTGAT  2100
     GTTATAAAAA ATGATGGAAT TTACTCGAGG TATTTTTTCT CCTTTGCTGC AAATGGTAGA  2160
     TATAGCTTGA AAGTGCATGT CAATCACTCT CCCAGCATAA GCACCCCAGC CCACTCTATT  2220
     CCAGGGAGTC ATGCTATGTA TGTACCAGGT TACACAGCAA ACGGTAAATAT TCAGATGAAT  2280
     GCTCCAAGGA AATCAGTAGG CAGAAATGAG GAGGAGCGAA AGTGGGGCTT TAGCCGAGTC  2340
     AGCTCAGGAG GCTCCTTTTC AGTGCTGGGA GTTCCAGCTG GCCCCCACCC TGATGTGTTT  2400
     CCACCATGCA AAATTATTGA CCTGGAAGCT GTAAAAGTAG AAGAGGAATT GACCCTATCT  2460
     TGGACAGCAC CTGGAGAAGA CTTTGATCAG GGCCAGGCTA CAAGCTATGA AATAAGAATG  2520
     AGTAAAAGTC TACAGAATAT CCAAGATGAC TTTAACAATG CTATTTTAGT AAATACATCA  2580
     AAGCGAAATC CTCAGCAAGC TGGCATCAGG GAGATATTTA CGTTCTCACC CCAGATTTCC  2640
     ACGAATGGAC CTGAACATCA GCCAAATGGA GAAACACATG AAAGCCACAG AATTTATGTT  2700
     GCAATACGAG CAATGGATAG GAACTCCTTA CAGTCTGCTG TATCTAACAT TGCCCAGGCG  2760
     CCTCTGTTTA TTCCCCCCAA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAAA  2820
     GGAGTTTTAA CAGCAATGGG TTTGATAGGA ATCATTTGCT TTATTATAGT TGTGACACAT  2880
     CATACTTTAA GCAGGAAAAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATTATAA  2940
     ATAAATATCC AAAGTGTCTT CCTTCTTAGA TATAAGACCC ATGGCCTTCG ACTACAAAAA  3000
     CATACTAACA AAGTCAAATT AACATCAAAA CTGTATTAAA ATGCATTGAG TTTTTGTACA  3060
     ATACAGATAA GATTTTTACA TGGTAGATCA ACAATTCTTT TTGGGGGTAG ATTAGAAAAC  3120
     CCTTACACTT TGGCTATGAA CAAATAATAA AAATTATTCT TTAAAGTAAT GTCTTTAAAG  3180
     GCAAAGGAA GGGTAAAGTC GGACCAGTGT CAAGGAAAGT TTGTTTTATT GAGGTGGAAA  3240
     AATAGCCCCA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAACTGTCTG TGTGAAGCAA  3300
     TCATTTAGTT ACTTTGATTA ATTTTTCTTT TCTCCTTATC TGTGCAGTAC AGGTTGCTTG  3360
     TTTACATGAA GATCATGCTA TATTTTATAT ATGTAGCCCC TAATGCAAAG CTCTTTACCT  3420
     CTTGCTATTT TGTTATATAT ATTTCAGATG ACATCTCCCT GCTAATGCTC AGAGATCTTT  3480
     TTTCACTGTA AGAGGTAACC TTTAACAATA TGGGTATTAC CTTTGTCTCT TCATACCGGT  3540
     TTTATGACAA AGGTCTATTG AATTTATTTG TNTGTAAGTT TCTACTCCCA TCAAAGCAGC  3600
     TTTCTAAGTT TATTGCCTTG GGTTATTATG GAATGATAGT TATAGCCCCN TATAATGCCT  3660
     TACCTAGGAA A
```

Seq ID NO: 318 Protein sequence
Protein Accession #: NP_006527.1

```
     1         11        21        31        41        51
     |         |         |         |         |         |
     MTQRSIAGPI CNLKFVTLLV ALSSELPFLG AGVQLQDNGY NGLLIAINPQ VPENQNLISN    60
```

```
            IKEMITEASF  YLFNATKRRV  FFRNIKILIP  ATWKANNNSK  IKQESYEKAN  VIVTDWYGAH    120
            GDDPYTLQYR  GCGKEGKYIH  FTPNFLLNDN  LTAGYGSRGR  VFVHEWAHLR  WGVFDEYNND    180
            KPFYINGQNQ  IKVTRCSSDI  TGIFVCEKGP  CPQENCIISK  LFKEGCTFIY  NSTQNATASI    240
            MFMQSLSSVV  EFCNASTHNQ  EAPNLQNQMC  SLRSAWDVIT  DSADFHHSFP  MNGTELPPPP    300
            TFSLVQAGDK  VVCLVLDVSS  KMAEADRLLQ  LQQAAEFYLM  QIVEIHTFVG  IASFDSKGEI    360
            RAQLHQINSN  DDRKLLVSYL  PTTVSAKTDI  SICSGLKKGF  EVVEKLNGKA  YGSVMILVTS    420
            GDDKLLGNCL  PTVLSSGSTI  HSIALGSSAA  PNLEBLSRLT  GGLKPFVPDI  SNSNSMIDAF    480
            SRISSGTGDI  FQQHIQLEST  GENVKPHHQL  KNTVTVDNTV  GNDTMPLVIW  QASGPPEIIL    540
            FDPDGRKYYT  NNFITNLTFR  TASLWIPGTA  KPGHWTYTLN  NTHHSLQALK  VTVTSRASNS    600
            AVPPATVEAF  VERDSLHFPH  PVMIYANVKQ  GFYPILNATV  TATVEPETGD  PVTLRLLDDG    660
            AGADVIKNDG  IYSRYFPSPA  ANGRYSLKVH  VNHSPSISTP  AHSIFGSHAM  YVPGYTANGN    720
            IQMNAPRKSV  GRNEEERKWG  FSRVSSGGSF  SVLGVPAGPH  PDVFPPCKII  DLEAVKVEEE    780
            LTLSWTAPGE  DFDQGQATSY  EIRMSKSLQN  IQDDFNNAIL  VNTSKRNPQQ  AGIREIFTFS    840
            PQISTNGPEH  QPNGETHESH  RIYVAIRAMD  RNSLQSAVSN  IAQAPLFIPP  NSDPVPARDY    900
            LILKGVLTAM  GLIGIICLII  VVTHHTLSRK  KRADKKENGT  KLL
```

Seq ID NO: 319 DNA sequence
Nucleic Acid Accession #: NM_000228.1
Coding sequence: 82..3600

```
            1          11         21         31         41         51
            |          |          |          |          |          |
            GCTTTCAGGC GATCTGGAGA AAGAACGGCA GAACACACAG CAAGGAAAGG TCCTTTCTGG     60
            GGATCACCCC ATTGGCTGAA GATGAGACCA TTCTTCCTCT TGTGTTTTGC CCTGCCTGGC    120
            CTCCTGCATG CCCAACAAGC CTGCTCCCGT GGGGCCTTGG ATCCACCTGT TGGGGACCTG    180
            CTTGTTGGGA GGACCCGGTT TCTCCGAGCT TCATCTACCT GTGGACTGAC CAAGCCTGAG    240
            ACCTACTGCA CCCAGTATGG CGAGTGGCAG ATGAAATGCT GCAAGTGTGA CTCCAGGCAG    300
            CCTCACAACT ACTACAGTCA CCGAGTAGAG AATGTGGCTT CATCCTCCGG CCCCATGCGC    360
            TGGTGGCAGT CCCAGAATGA TGTGAACCCT GTCTCTCTGC AGCTGGACCT GGACAGGAGA    420
            TTCCAGCTTC AAGAAGTCAT GATGGAGTTC CAGGGGCCCA TGCCCGCCGG CATGCTGATT    480
            GAGCGCTCCT CAGACTTCGG TAAGACCTGG CGAGTGTACC AGTACCTGGC TGCCGACTGC    540
            ACCTCCACCT TCCCTCGGGT CCGCCAGGGT CGGCCTCAGA GCTGGCAGGA TGTTCGGTGC    600
            CAGTCCCTGC CTCAGAGGCC TAATGCACGC CTAAATGGGG GGAAGGTCCA ACTTAACCTT    660
            ATGGATTTAG TGTCTGGGAT TCCAGCAACT CAAAGTCAAA AAATTCAAGA GGTGGGGGAG    720
            ATCACAAACT TGAGAGTCAA TTTCACCAGG CTGGCCCCTG TGCCCCAAAG GGGCTACCAC    780
            CCTCCCAGCG CCTACTATGC TGTGTCCCAG CTCCGTCTGC AGGGGAGCTG CTTCTGTCAC    840
            GGCCATGCTG ATCGCTGCGC ACCCAAGCCT GGGGCCTCTG CAGGCCCCTC CACCGCTGTG    900
            CAGGTCCACG ATGTCTGTGT CTGCCAGCAC AACACTGCCG GCCCAAATTG TGAGCGCTGT    960
            GCACCCTTCT ACAACAACCG GCCCTGGAGA CCGGCGGAGG GCCAGGACGC CCATGAATGC   1020
            CAAAGGTGCG ACTGCAATGG GCACTCAGAG ACATGTCACT TTGACCCCGC TGTGTTTGCC   1080
            GCCAGCCAGG GGGCATATGG AGGTGTGTGT GACAATTGCC GGGACCACAC CGAAGGCAAG   1140
            AACTGTGAGC GGTGTCAGCT GCACTATTTC CGGAACCGGC GCCCGGGAGC TTCCATTCAG   1200
            GAGACCTGCA TCTCCTGCGA GTGTGATCCG GATGGGGCAG TGCCAGGGGC TCCCTGTGAC   1260
            CCAGTGACCG GGCAGTGTGT GTGCAAGGAG CATGTGCAGG GAGAGCGCTG TGACCTATGC   1320
            AAGCCGGGCT TCACTGGACT CACCTACGCC AACCCGCAGG GCTGCCACCG CTGTGACTGC   1380
            AACATCCTGG GGTCCCGGAG GGACATGCCG TGTGACGAGG AGAGTGGGCG CTGCCTTTGT   1440
            CTGCCCAACG TGGTGGGTCC CAAATGTGAC CAGTGTGCTC CCTACCACTG GAAGCTGGCC   1500
            AGTGGCCAGG GCTGTGAACC GTGTGCCTGC GACCGGCACA ACTCCCCTCA GCCCACAGTG   1560
            CAACCAGTTC ACAGGGCAGT GCCCTGTCGG GAAGGCTTTG GTGGCCTGAT GTGCAGCGCT   1620
            GCAGCCATCC GCCAGTGTCC AGACCGGACC TATGGAGACG TGGCCACAGG ATGCCGAGCC   1680
            TGTGACTGTG ATTTCCGGGG AACAGAGGGC CCGGGCTGCG ACAAGGCATC AGGCCGCTGC   1740
            CTCTGCCGCC CTGGCTTGAC CGGGCCCCGC TGTGACCAGT GCCAGCGAGG CTACTGCAAT   1800
            CGCTACTGCG TGTGCGTGGC CTGCCACCCT TGCTTCCAGA CCTATGATGC GGACCTCCGG   1860
            GAGCAGGCCC TGCGCTTTGG TAGACTCCGC AATGCCACCG CCAGCCTGTG TCAGGGCCT    1920
            GGGCTGGAGG ACGGTGGCCT GGCCTCCCGG ATCCTAGATG CAAAGAGTAA GATTGAGCAG   1980
            ATCCGAGCAG TTCTCAGCAG CCCCGCAGTC ACAGAGCAGG AGGTGGCTCA GGTGGCCAGT   2040
            GCCATCCTCT CCCTCAGGCG AACTCTCCAG GGGCTGCAGC TGGATCTGCC CCTGGAGGAG   2100
            GAGACGTTGT CCCTTCCGAG AGACCTGGAG AGTCTTGACA GAAGCTTCAA TGGTCTCCTT   2160
            ACTATGTATC AGAGGAAGAG GGAGCAGTTT GAAAAAATAA GCAGTGCTGA TCCTTCAGGA   2220
            GCCTTCCGGA TGCTGAGCAC AGCCTACGAG CAGTCAGCCC AGGCTGCTCA GCAGGTCTCC   2280
            GACAGCTCGC GCCTTTTGGA CCAGCTCAGG GACAGCGGGA GAGAGGCAGA GAGGCTGGTG   2340
            CGGCAGGCGG GAGGAGGAGG AGGCACCGGC AGCCCCAAGC TTGTGGCCCT GAGGCTGGAG   2400
            ATGTCTTCGT TGCCTGACCT GACACCCACC TTCAACAAGC TCTGTGGCAA CTCCAGGCAG   2460
            ATGGCTGCA CCCCAATATC ATGCCCTGGT GAGCTATGTC CCCAAGACAA TGGCACAGCC   2520
            TGTGGCTCCC GCTGCAGGGG TGTCCTTCCC AGGGCCGGTG GGGCCTTCTT GATGCGGGGG   2580
            CAGGTGGCTG AGCAGCTGCG GGGCTTCAAT GCCCAGCTCC AGCGGACCAG GCAGATGATT   2640
            AGGGCAGCCG AGGAATCTGC CTCACAGATT CAATCCAGTG CCCAGCGCTT GGAGACCCAG   2700
            GTGAGCGCCA GCCGCTCCCA GATGGAGGAA GATGTCAGAC GCACACGGCT CCTAATCCAG   2760
            CAGGTCCGGG ACTTCCTAAC AGACCCCGAC ACTGATGCAG CCACTATCCA GGAGGTCAGC   2820
            GAGGCCGTGC TGGCCCTGTG GCTGCCCACA GACTCAGCTA CTGTTCTGCA GAAGATGAAT   2880
            GAGATCCAGG CCATTGCAGC CAGGCTCCCC AACGTGGACT TGGTGCTGTC CCAGACCAAG   2940
            CAGGACATTG CGGTGCGCCG CCGGTTGCAG GCTGAGGCTG AGGAAGCCGA GAGCCAGGCC   3000
            CATGCAGTGG AGGGCCAGGT GGAAGATGTG GTTGGGAACC TGCGGCAGGG GACAGTGGCA   3060
            CTGCAGGAAG CTCAGGACAC CATGCAAGGC ACCAGCCGCT CCCTTCGGCT TATCCAGGAC   3120
            AGGGTTGCTG AGGTTCAGCA GGTACTGCGG CCAGCAGAAA AGCTGGTGAC AAGCATGACC   3180
            AAGCAGCTGG GTGACTTCTG GACACGGATG GAGGAGCTCC GCCACCAAGC CGGCAGCAG    3240
            GGGGCAGAGG CAGTCCAGGC CCAGCAGCTT CGGGAAGGTG CCAGCGAGCA GGCATTGAGT   3300
            GCCCAAGAGG GATTTGAGAG AATAAAACAA AAGTATGCTG AGTTGAAGGA CCGGTTGGGT   3360
            CAGAGTTCCA TGCTGGGTGA GCAGGGTGCC CGGATCCAGA GTGTGAAGAC AGAGGCAGAG   3420
            GAGCTGTTTG GGGAGACCAT GGAGATGATG GACAGGATGA AAGACATGGA GTTGGAGCTG   3480
            CTGCGGGGCA GCCAGGCCAT CATGCTGCGC TCGGCGGACA TGACAGGACT GGAGAAGCGT   3540
            GTGGAGCAGA TCCGTGACCA CATCAATGGG CGCGTGCTCT ACTATGCCAC CTGCAAGTGA   3600
            TGCTACAGCT TCCAGCCCGT TGCCCCACTC ATCTGCCGCC TTTGCTTTTG GTTGGGGGCA   3660
            GATTGGGTTG GAATGCTTTC CATCTCCAGG AGACTTTCAT GCAGCCTAAA GTACAGCCTG   3720
```

```
GACCACCCCT  GGTGTGTAGC  TAGTAAGATT  ACCCTGAGCT  GCAGCTGAGC  CTGAGCCAAT  3780
GGGACAGTTA  CACTTGACAG  ACAAAGATGG  TGGAGATTGG  CATGCCATTG  AAACTAABAG  3840
CTCTCAAGTC  AAGGAAGCTG  GGCTGGGCAG  TATCCCCCGC  CTTTAGTTCT  CCACTGGGGA  3900
GGAATCCTGG  ACCAAGCACA  AAAACTTAAC  AAAAGTGATG  TAAAAATGAA  AAGCCAAATA  3960
AAAATCTTTG  G
```

Seq ID NO: 320 <u>Protein sequence</u>
Protein Accession #: NP_000219.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MRPPFLLCPA  LPGLLHAQQA  CSRGACYPPV  GDLLVGRTRF  LRASSTCGLT  KPETYCTQYG  60
EWQMKCCKCD  SRQPHNYYSH  RVENVASSSG  PMRWWQSQND  VNPVSLQLDL  DRRFQLQEVM  120
MEPQGPMPAG  MLIERSSDFG  KTWRVYQYLA  ADCTSTFPRV  RQGRPQSNQD  VRCQSLPQRP  180
NARLNGGKVQ  LNLMDLVSGI  PATQSQKIQE  VGEITNLRVN  PTRLAPVPQR  GYHPPSAYYA  240
VSQLRLQGSC  FCHGHADRCA  PKPGASAGPS  TAVQVHDVCV  CQHNTAGPNC  ERCAPFYNNR  300
PWRPAEGQDA  HECQRCDCNG  HSETCHFDPA  VFAASQGAYG  GVCDNCRDHT  EGKNCERCQL  360
HYFRNRRPGA  SIQETCISCE  CDPDGAVPGA  PCDPVTGQCV  CKEHVQGERC  DLCKPGFTGL  420
TYANPQGCHR  CDCNILGSRR  DMPCDEESGR  CLCLPNVVGP  KCDQCAPYHW  KLASGQGCEP  480
CACDPHNSPQ  PTVQPVHRAV  PCREGFGGLM  CSAAAIRQCP  DRTYGDVATG  CRACDCDFRG  540
TEGPGCDKAS  GRCLCRPGLT  GPRCDQCQRG  YCNRYPVCVA  CHPCFQTYDA  DLREQALRFG  600
RLRNATASLW  SGPGLEDRGL  ASRILDAKSK  IEQIRAVLSS  PAVTSQEVAQ  VASAILSLRR  660
TLQGLQLDLP  LEEETLSLPR  DLESLDRSFN  GLLTMYQRKR  EQFEKISSAD  PSGAFRMLST  720
AYEQSAQAAQ  QVSDSSRLLD  QLRDSRREAE  RLVRQAGGGG  GTGSPKLVAL  RLEMSSLFDL  780
TPTFNKLCGN  SRQMACTPIS  CPGELCPQDN  GTACGSRCRG  VLPRAGGAFL  MAGQVAEQLR  840
GFNAQLQRTR  QMIRAAEESA  SQIQSSAQRL  ETQVSASREQ  MREDVRRTRL  LIQQVRDFLT  900
DPDTDAATIQ  EVSEAVLALW  LPTDSATVLQ  KMNEIQAIAA  RLPNVDLVLS  QTKQDIARAR  960
RLQAEAEBAR  SRAHAVEGQV  EDVVGNLRQG  TVALQEAQDT  MQGTSRSLRL  IQDRVAEVQQ  1020
VLRPAEKLVT  SMTKQLGDFW  TRMEELRHQA  RQQGAEAVQA  QQLAEGASEQ  ALSAQEGFER  1080
IKQKYAELKD  RLGQSSMLGE  QGARIQSVKT  EAEELFGETM  EMMDRMQDME  LELLRGSQAI  1140
MLRSADLTGL  EKRVEQIRDH  INGRVLYYAT  CK
```

Seq ID NO: 321 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_001944.1
Coding sequence: 84..3083

```
1          11         21         31         41         51
|          |          |          |          |          |
TTTTCTTAGA  CATTAACTGC  AGACGGCTGG  CAGGATAGAA  GCAGCGGCTC  ACTTGGACTT  60
TTTCACCAGG  GAAATCAGAG  ACAATGATGG  GGCTCTTCGC  CAGAACTACA  GGGGCGTCTG  120
CCATCTTCGT  GGTGGTCATA  TTGGTTCATG  GAGAATTGCG  AATAGAGACT  AAAGGTCAAT  180
ATGATGAAGA  AGAGATGACT  ATGCAACAAG  CTAAAAGAAG  GCAAAAACGT  GAATGGGTGA  240
AATTTGCCAA  ACCCTGCAGA  GAAGGAGAAG  ATAACTCAAA  AAGAAACCCA  ATTGCCAAGA  300
TTACTTCAGA  TTACCAAGCA  ACCCAGAAAA  TCACCTACCG  AATCTCTGGA  GTGGGAATCG  360
ATCAGCCGCC  TTTTGGAATC  TTTGTTGTTG  ACAAAAACAC  TGGAGATATT  AACATAACAG  420
CTATAGTCGA  CCGGGAGGGA  ACTCCAAGCT  TCCTGATCAC  ATGTCGGGCT  CTAAATGCCC  480
AAGGACTAGA  TGTAGAGAAA  CCACTTATAC  TAACGGTTAA  AATTTTGGAT  ATTAATGATA  540
ATCCTCCAGT  ATTTTCACAA  CAAATTTTCA  TGGGTGAAAT  TGAAGAAAAT  AGTGCCTCAA  600
ACTCACTGGT  GATGATACTA  AATGCCACAG  ATGCAGATGA  ACCCAAACCAC  TTGAATTCTA  660
AAATTGCCTT  CAAAATTGTC  TCTCAGGAAC  CAGCAGGCAC  ACCCATGTTC  CTCCTAAGCA  720
GAAACACTGG  GGAAGTCCGT  ACTTTGACCA  ATTCTCTTGA  CCGAGAGCAA  GCTAGCAGCT  780
ATCGTCTGGT  TGTGAGTGGT  GCAGACAAAG  ATGGAGAGG  ACTATCAACT  CAATGTGAAT  840
GTAATATTAA  AGTGAAAGAT  GTCAACGATA  ACTTCCCAAT  GTTTAGAGAC  TCTCAGTATT  900
CAGCACGTAT  TGAAGAAAAT  ATTTTTAAGTT  CTGAATTACT  TCGATTTCAA  GTAACAGATT  960
TGGATGAAGA  GTACACAGAT  AATTGGCTTG  CAGTATATTT  CTTTACCTCT  GGGAATGAAG  1020
GAAATTGGTT  TGAAATACAA  ACTGATCCTA  GAACTAATGA  AGGCATCCTA  AAAGTGGTGA  1080
AGGCTCTAGA  TTATGAACAA  CTACAAAGCG  TGAAACTTAG  TATTGCTGTC  AAAAACAAAG  1140
CTGAATTTCA  CCAATCAGTT  ATCTCTCGAT  ACCGAGTTCA  GTCAACCCCA  GTCACAATTC  1200
AGGTAATAAA  TGTAAGAGAA  GGAATTGCAT  TCCGTCCTGC  TTCCAAGACA  TTTACTGTGC  1260
AAAAAGGCAT  AAGTAGCAAA  AAATTGGTGG  ATTATATCCT  GGGAACATAT  CAAGCCATCG  1320
ATGAGGACAC  TAACAAAGCT  GCCTCAAATG  TCAAATATGT  CATGGGACGT  AACGATGGTG  1380
GATACCTAAT  GATTGATTCA  AAAACTGCTG  AAATCCAATT  TGTCAAAAAT  ATGAACCGGA  1440
ATTCTACTTT  CATAGTTAAC  AAAACAATCA  CAGCTGAGGT  TCTGGCCATA  GATGAATACA  1500
CGGGTAAAAC  TTCTACAGGC  ACGGTATATG  TTAGAGTACC  CGATTCAAT  GACAATTGTC  1560
CAACAGCTGT  CCTCGAAAAA  GATGCAGTTT  GCAGTTCTTC  ACCTTCCGTG  GTTGTCTCCG  1620
CTAGAACACT  GAATAATAGA  TACACTGGCC  CCTATACATT  TGCACTGGAA  GATCAACCTG  1680
TAAAGTTGCC  TGCCGTATGG  AGTATCACAA  CCCTCAATGC  TACCTCGGCC  CTCCTCAGAG  1740
CCCAGGAACA  GATACCTCCT  GGAGTATACC  ACATCTCCCT  GGTACTTACA  GACAGTCAGA  1800
ACAATCGGTG  TGAGATGCCA  CGCAGCTTGA  CACTGGAAGT  CTGTCAGTGT  GACAACCAGG  1860
GCATCTGTGG  AACTTCTTAC  CCAACCACA  GCCCTGGGAC  CAGGTATGGC  AGGCCGCACT  1920
CAGGGAGGCT  GGGGCCTGCC  GCCATCGGCC  TGCTGCTCCT  TGGTCTCCTG  CTGCTGCTGT  1980
TGGCCCCCCT  TCTGCTGTTG  ACCTGTGACT  GTGGGGCAGG  TTCTACTGGG  GGAGTGACAG  2040
GTGGTTTTAT  CCCAGTTCCT  GATGGCTCAG  AAGGAACAAT  TCATCAGTGG  GGAATTGAAG  2100
GAGCCCATCC  TGAAGACAAG  GAAATCACAA  ATATTTGTGT  GCCTCCTGTA  ACAGCCAATG  2160
GAGCCGATTT  CATGGAAAGT  TCTGAAGTTT  GTACAAATAC  GTATGCCAGA  GGCACAGCCG  2220
TGGAAGGCAC  TTCAGGAATG  GAAATGACCA  CTAAGCTTGG  AGCAGCCACT  GAATCTGGAG  2280
GTGCTGCAGG  CTTTGCAACA  GGGACAGTGT  CAGGAGCTGC  TTCAGGATTC  GGAGCAGCCA  2340
CTGGAGTTGG  CATCTGTTCC  TCAGGGCAGT  CTGGAACCAT  GAGAACAAGG  CATTCCACTG  2400
GAGGAACCAA  TAAGGACTAC  GCTGATGGGG  CGATAAGCAT  GAATTTTCTG  GACTCCTACT  2460
TTTCTCAGAA  AGCATTTGCC  TGTGCGGAGG  AAGACGATGG  CCAGGAAGCA  AATGACTGCT  2520
TGTTGATCTA  TGATAATGAA  GGCGCAGATG  CCACTGGTTC  TCCTGTGGGC  TCCGTGGGTT  2580
GTTGCAGTTT  TATTGCTGAT  GACCTGGATG  ACAGCTTCTT  GGACTCACTT  GGACCCAAAT  2640
TTAAAAAACT  TGCAGAGATA  AGCCTTGGTG  TTGATGGTGA  AGGCAAAGAA  GTTCAGCCAC  2700
CCTCTAAAGA  CAGCGGTTAT  GGGATTGAAT  CCTGTGGCCA  TCCCATAGAA  GTCCAGCAGA  2760
```

```
CAGGATTTGT TAAGTGCCAG ACTTTGTCAG GAAGTCAAGG AGCTTCTGCT TTGTCCGCCT  2820
CTGGGTCTGT CCAGCCAGCT GTTTCCATCC CTGACCCTCT GCAGCATGGT AACTATTTAG  2880
TAACGGAGAC TTACTCGGCT TCTGGTTCCC TCGTGCAACC TTCCACTGCA GGCTTTGATC  2940
CACTTCTCAC ACAAAATGTG ATAGTGACAG AAAGGGTGAT CTGTCCCATT TCCAGTGTTC  3000
CTGGCAACCT AGCTGGCCCA ACGCAGCTAC GAGGGTCACA TACTATGCTC TGTACAGAGG  3060
ATCCTTGCTC CCGTCTAATA TGACCAGAAT GAGCTGGAAT ACCACACTGA CCAAATCTGG  3120
ATCTTTGGAC TAAAGTATTC AAAAATAGCAT AGCAAAGCTC ACTGTATTGG GCTAATAATT  3180
TGGCACTTAT TAGCTTCTCT CATAAACTGA TCACGATTAT AAATTAAATG TTTGGGTTCA  3240
TACCCCAAAA GCAATATGTT GTCACTCCTA ATTCTCAAGT ACTATTCAAA TTGTAGTAAA  3300
TCTTAAAGTT TTTCAAAACC CTAAAATCAT ATTCGC
```

Seq ID NO: 322 Protein sequence
Protein Accession #: NP_001935.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MMGLFPRTTG ALAIPVVVIL VHGELRIETK GQYDEEEMTM QQAKRRQKRE WVKFAKPCRE   60
GSDNSKRNPI AKITSDYQAT QKITYRISGV GIDQPPFGIF VVDKNTGDIN ITAIVDREET  120
PSFLITCRAL NAQGLDVEKP LILTVKILDI NDNPPVFSQQ IFMGEIEENS ASNSLVMILN  180
ATDADEPNHL NSKIAFKIVS QEPAGTPMFL LSRNTGEVRT LTNSLDREQA SSYRLVVSGA  240
DKDGEGLSTQ CECNIKVKDV NDNFPMFRDS QYSARIEENI LSSELLRFQV TDLDEEYTDN  300
WLAVYFFTSG NEGNWFEIQT DPRTNEGILK VVKALDYEQL QSVKLSIAVK NKAEFHQSVI  360
SRYRVQSTPV TIQVINVREG IAFRPASKTF TVQKGISSKK LVDYILGTYQ AIDEDTNKAA  420
SNVKYVMGRN DGGYLMIDSK TAEIKFVKNM NRDSTFIVNK TITAEVLAID EYTGKTSTGT  480
VYVRVPDFND NCPTAVLEKD AVCSSSPSVV VSARTLNNRY TGPYTFALED QPVKLPAVWS  540
ITTLNATSAL LRAQSQIPPG VYHISLVLTD SQNNRCEMPR SLTLEVCQCD NRGICGTSYP  600
TTSPGTRYGR PHSGRLGPAA IGLLILGLLL LLLAPLLLLT CDCGAGSTGG VTGGFIPVPD  660
GSEGTIHQWG IEGAHPEDKE ITNICVPPVT ANGADFMESS EVCTNTYARG TAVEGTSGME  720
MTTKLGAATE SGGAAGFATG TVSGAASGFG AATGVGICSS GQSGTMRTRH STGGTNKDYA  780
DGAISMNFLD SYFSQKAPAC AEEDDGQEAN DCLLIYDNEG ADATGSPVGS VGCCSFIADD  840
LDDSFLDSLG PKFKKLAEIS LGVDGEGKEV QPPSKDSGYG IESCGHPIEV QQTGFVKCQT  900
LSGSQGASAL SASGSVQPAV SIPDPLQHGN YLVTETYSAS GSLVQPSTAG FDPLLTQNVI  960
VTERVICPIS SVPGNLAGPT QLRGSHTMLC TEDPCSRLI
```

Seq ID NO: 323 DNA sequence
Nucleic Acid Accession #: XM_058069.2
Coding sequence: 1..1413

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAAGTTTC TTCTAATACT GCTCCTGCAG GCCACTGCTT CTGGAGCTCT TCCCCTGAAC   60
AGCTCTACAA GCCTGGAAAA AAATAATGTG CTATTTGGTG AAAGATACTT AGAAAAATTT  120
TATGGCCTTG AGATAAACAA ACTTCCAGTG ACAAAAATGA AATATAGTGG AAACTTAATG  180
AAGGAAAAAA TCCAAGAAAT GCAGCACTTC TTGGGTCTGA AAGTGACCGG GCAACTGGAC  240
ACATCTACCC TGGAGATGAT GCACGCACCT CGATGTGGAG TCCCCGATGT CCATCATTTC  300
AGGGAAATGC CAGGGGGGCC CGTATGGAGG AAACATTATA TCACCTACAG AATCAATAAT  360
TACACACCTG ACATGAACCG TGAGGATGTT GACTACGCAA TCCGGAAAGC TTTCCAAGTA  420
TGGAGTAATG TTACCCCCTT GAAATTCAGC AAGATTAACA CAGGCATGGC TGACATTTTG  480
GTGGTTTTTG CCCGTGGAGC TCATGGAGAC TTCCATGCTT TTGATGGCAA AGGTGGAATC  540
CTAGCCCATG CTTTTGACCC TGGATCTGGC ATTGGAGGGG ATGCACATTT CGATGAGGAC  600
GAATTCTGGA CTACACATTC AGGAGGCACA AACTTGTTCC TCACTGCTGT TCACGAGATT  660
GGCCATTCCT TAGGTCTTGG CCATTCTAGT GATCCAAAGG CCGTAATGTT CCCCACCTAC  720
AAATATGTTG ACATCAACAC ATTTCGCCTC TCTGCTGATG ACATACGTGG CATTCAGTCC  780
CTGTATGGAG ACCCAAAAGA GAACCAACGC TTGCCAAATC CTGACAATTC AGAACCAGCT  840
CTCTGTGACC CCAATTTGAG TTTTGATGCT GTCACTACCG TGGGAAATAA GATCTTTTTC  900
TTCAAAGACA GGTTCTTCTG GCTGAAGGTT TCTGAGACGA CAAAGACCAG TGTTAATTTA  960
ATTTCTTCCT TATGGCCAAC CTTGCCATCT GGCATTGAAG CTGCTTATGA AATTGAAGCC  1020
AGAAATCAAG TTTTTCTTTT TAAAGATGAC AAATACTGGT AATTAGCAA TTTAAGACCA  1080
GAGCCAAATT ATCCCAAGAG CATACATTCT TTTGGTTTTC CTAACTTTGT GAAAAAAATT  1140
GATGCAGCTG TTTTTAACCC ACGTTTTTAT AGGACCTACT TCTTTGTAGA TAACCAGTAT  1200
TGGAGGTATG ATGAAAGGAG ACAGATGATG GACCCTGGTT ATCCCAAACT GATTACCAAG  1260
AACTTCCAAG GAATCGGGCC TAAAATTGAT GCAGTCTTCT ACTCTAAAAA CAAATACTAC  1320
TATTTCTTCC AAGGATCTAA CCAATTTGAA TATGACTTCC TACTCCAACG TATCACCAAA  1380
ACACTGAAAA GCAATAGCTG GTTTGGTTGT TGA
```

Seq ID NO: 324 Protein sequence
Protein Accession #: P39900

```
1          11         21         31         41         51
|          |          |          |          |          |
MKFLLILLLQ ATASGALPLN SSTSLEKNNV LFGERYLEKF YGLEINKLPV TKMKYSGNLM   60
KEKIQEMQHF LGLKVTGQLD TSTLEMMHAP RCGVPDVHHF REMPGGFVWR KHYITYRINN  120
YTPDMNREDV DYAIRKAFQV WSNVTPLKFS KINTGMADIL VVFARGAHGD FHAFDGKGGI  180
LAHAFGPGSG IGGDAHFDED EFWTTHSGGT NLFLTAVHEI GHSLGLGHSS DPKAVMFPTY  240
KYVDINTFRL SADDIRGIQS LYGDPKENQR LPNPDNSEPA LCDPNLSFDA VTTVGNKIFF  300
FKDRFFWLKV SERPKTSVNL ISSLWPTLPS GIEAAYEIEA RNQVFLFKDD KYWLISNLRP  360
EPNYPKSIHS FGFPNFVKKI DAAVFNPRFY RTYFFVDNQY WRYDERRQMM DPGYPKLITK  420
NFQGIGPKID AVFYSKNKYY YFFQGSNQFE YDFLLQRITK TLKSNSWFGC
```

Seq ID NO: 325 DNA sequence
Nucleic Acid Accession #: NM_024423.1
Coding sequence: 64..2590

```
1          11         21         31         41         51
```

```
            |           |           |           |           |           |
          GGCAGGTCTC GCTCTCGGCA CCCTCCCGGC GCCCGCGTTC TCCTGGCCCT GCCCGGCATC    60
          CCGATGGCCG CCGCTGGGCC CCGGCGCTCC GTGCGCGGAG CCGTCTGCCT GCATCTGCTG   120
          CTGACCCTCG TGATCTTCAG TCGTGATGGT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA   180
          CCTTCTAAAC TAGAGGCAGA CAAAAATAATT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG   240
          TCTGCAGACC TCATCCGGTC AAGTGATCCT GATTTCAGAG TTCTAAATGA TGGGTCAGTG   300
          TACACAGCCA GGGCTGTTGC GCTGTCTGAT AAGAAAAGAT CATTTACCAT ATGGCTTTCT   360
          GACAAAAGGA AACAGACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA   420
          TCGAAGACAA GACACACTAG AGAAACTGTT CTCAGGCGTG CCAAGAGGAG ATGGGCACCT   480
          ATTCCTTGCT CTATGCAAGA GAATTCCTTG GGCCCTTTCC CATTGTTTCT TCAACAAGTT   540
          GAATCTGATG CAGCACACGAA CTATACTGTC TTCTACTCAA TAAGTGGACG TGGAGTTGAT   600
          AAAGAACCTT TAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTATT TTGCACTCGG   660
          CCTGTGGATC GTGAAGAATA TGATGTTTTT GATTTGATTG CTTATGCGTC AACTGCAGAT   720
          GGATATTCAG CAGATCTGCC CCTCCCACTA CCCATCAGGG TAGAGGATGA AAATGACAAC   780
          CACCCTGTTT TCACAGAAGC AATTTATAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT   840
          ACTACAGTGG GGGTGGTTTG TGCCACAGAC AGAGATGAAC CGGACACAAT GCATACGCGC   900
          CTGAAATACA GCATTTTGCA GCAGACACCA AGGTCACCTG GGCTCTTTTC TGTGCATCCC   960
          AGCACAGGCG TAATCACCAC AGTCTCTCAT TATTTGGACA GAGAGGTTGT AGACAAGTAC  1020
          TCATTGATAA TGAAAGTACA AGACATGGAT GGCCAGTTTT TTGGATTGAT AGGCACATCA  1080
          ACTTGTATCA TAACAGTAAC AGATTCAAAT GATAATGCAC CCACTTTCAG ACAAAATGCT  1140
          TATGAAGCAT TTGTAGAGGA AAATGCATTC AATGTGGAAA TCTTACGAAT ACCTATAGAA  1200
          GATAGGGATT TAATTAACAC TGCCAATTGG AGAGTCAATT TTACCATTTT AAAGGGAAAT  1260
          GAAAATGGAC ATTTCAAAAT CAGCACAGAC AAAGAAACTA ATGAAGGTGT TCTTTCTGTT  1320
          GTAAAGCCAC TGAATTATGA AGAAAACCGT CAAGTGACC TGGAAATTGG AGTAAACAAT  1380
          GAAGCGCCAT TTGCTAGAGA TATTCCCAGA GTGACAGCCT TGAACAGAGC CTTGGTTACA  1440
          GTTCATGTGA GGGATCTGGA TGAGGGGCCT GAATGCACTC CTGCAGCCCA ATATGTGCGG  1500
          ATTAAAGAAA ACTTAGCAGT GGGGTCAAAG ATCAACGGCT ATAAGGCATA TGACCCCGAA  1560
          AATAGAAATG GCAATGGTTT AAGGTACAAA AAATTGCATG ATCCTAAAGG TTGGATCACC  1620
          ATTGATGAAA TTTCAGGGTC AATCATAACT TCCAAAATCC TGGATAGGGA GGTTGAAACT  1680
          CCCAAAAATG AGTTGTATAA TATTACAGTC CTGGCAATAG ACAAAGATGA TAGATCATGT  1740
          ACTGGAACAC TTGCTGTGAA CATTGAAGAT GTAAATGATA ATCCACCAGA AATACTTCAA  1800
          GAATATGTAG TCATTTGCAA ACCAAAAATG GGGTATACCG ACATTTTAGC TGTTGATCCT  1860
          GATGAACCTG TCCATGGAGC TCCATTTTAT TTCAGTTTGC CCAATACTTC TCCAGAAATC  1920
          AGTAGACTGT GGAGCCTCAC CAAAGTTAAT GATACAGCTG CCCGTCTTTC ATATCAGAAA  1980
          AATGCTGGAT TTCAAGAATA TACCATTCCT ATTACTGTAA AAGACAGGGC CGGCCAAGCT  2040
          GCAACAAAAT TATTGAGAGT TAATCTGTGT GAATGTACTC ATCCAACTCA GTGTCGTGCG  2100
          ACTTCAAGGA GTACAGGGAGT AATACTTGGA AAATGGGCAA TCCTTGCAAT ATTACTGGGT  2160
          ATAGCACTGC TCTTTTCTGT ATTGCTAACT TTAGTATGTG GAGTTTTTGG TGCAACTAAA  2220
          GGGAAACGTT TTCCTGAAGA TTTAGCACAG CAAAACTAA TTATATCAAA CACAGAAGCA  2280
          CCTGGAGACG ATAGAGTGTG CTCTGCCAAT GGATTTATGA CCCAAACTAC CAACAACTCT  2340
          AGCCAAGGTT TTTGTGGTAC TATGGGATCA GGAATGAAAA ATGGAGGGCA GGAAACCATT  2400
          GAAATGATGA AAGGAGGAAA CCAGACCTTG GAATCCTGCC GGGGGGCTGG GCATCATCAT  2460
          ACCCTGGACT CCTGCAGGGG AGGACACACG GAGGTGGACA ACTGCAGATA CACTTACTCG  2520
          GAGTGGCACA GTTTTACTCA ACCCCGTCTC GGTGAAGAAT CCATTAGAGG ACACACTGGT  2580
          TAAAAATTAA ACATAAAAGA AATTGCATCG ATGTAATCAG AATGAAGACC GCATGCCATC  2640
          CCAAGATTAT GTCCTCACTT ATAACTATGA GGGAAGAGGA TCTCCAGCTG GTTCTGTGGG  2700
          CTGCTGACTT GAAAAGCAGG AAGAAGATGG CCTTGACTTT TTAAATAATT TGGAACCCAA  2760
          ATTTATTACA TTAGCAGAAG CATGCACAAA GAGATAATGT CACAGTGCTA CAATTAGGTC  2820
          TTTGTCGACAC ATTCTGGAGG TTTCCAAAAA TAATATTGTA AAGTTCAATT TCAACATGTA  2880
          TGTATATGAT GATTTTTTTC TCAATTTTGA ATTATGCTAC TCACCAATTT ATATTTTTAA  2940
          AGCCAGTTGT TGCTTATCTT TTCCAAAAAG TGAAAAATGT TAAAACAGAC AACTGGTAAA  3000
          TCTCAAACTC CAGCACTGGA ATTAAGGTCT CTAAAGCATC TGCTCTTTTT TTTTTTTACG  3060
          GATATTTTAG TAATAAAATAT GCTGGATAAA TATTAGTCCA ACAATAGCTA AGTATAAAGAA  3120
          ATATCACATT ATTATGTATT CACTTTAAGT GATAGTTTAA AAAATAAACA AGAAATATTG  3180
          AGTATCACTA TGTGAAGAAA GTTTTGGAAA AGAAACAATG AAGACTGAAT TAAATTAAAA  3240
          ATGTTGCAGC TCATAAAGAA TTGGGACTCA CCCCTACTGC ACTACCAAAT TCATTTGACT  3300
          TTGGAGGCAA AATGTGTTGA AGTGCCCTAT GAAGTAGCAA TTTTCTATAG GAATATAGTT  3360
          GGAAATAAAT GTGTGTGTGT ATATTATTAT TAATCAATGC AATATTTAAA ATGAAATGAG  3420
          AACAAAGAGA AAAATGGTAA AAACTTGAAA TGAGGCTGGG GTATAGTTTG TCCTACAATA  3480
          GAAAAAAGAG AGAGCTTCCT AGGCCTGGGC TCTTAAATGC TGCATTATAA CTGAGTCTAT  3540
          GAGGAAATAG TTCCTGTCCA ATTTGTGTAA TTTGTTTAAA ATTGTAAATA AATTAAACTT  3600
          TTCTGGTTTC TGTGGGAAGG AAATAGGGAA TCCAATGGAA CAGTAGCTTT GCTTTGCAGT  3660
          CTGTTTCAAG ATTTCTGCAT CCACAAGTTA GTAGCAAACT GGGGAATACT CGCTGCAGCT  3720
          GGGGTTCCCT GCTTTTTGGT AGCAAGGGTC CAGAGATGAG GTGTTTTTTT CGGGGAGCTA  3780
          ATAACAAAAA CATTTTAAAA CTTACCTTTA CTGAAGTTAA ATCCTCTATT GCTGTTTCTA  3840
          TTCTCTCTTA TAGTGACCAA CATCTTTTTA ATTTAGATCG AAATAACCAT GTCCTCCTAG  3900
          AGTTTAGAGG CTAGAGGGAG CTGAGGGGAG GATCTTACTG AAAGCACCCT GGGGAGATTG  3960
          ATTGTCCTTA AACCTAAGCC CCACAAACTT GACACCTGAT CAGGTCTGGG AGCTACAAAA  4020
          TTTCATTTTT CTCCTCACTG CCCTTCTTCT GAGTGGCATT GGCCTGAATC AAGGAAAGCC  4080
          AGGCCTTGTG GGCCCCCTTC TTTGGCTTTT CTGCTAAAGC AACACCTCCA GCAGAGATTC  4140
          CCTTAAGTGA CTCCAGGTTT TCCACCATCC TTCAGCGTGA ATTAATTTTT AATCAGTTTG  4200
          CTTTCTCCAG AGAAATTTTA AAATAATAGA AGAAATAGAA ATTTGAATG TATAAAAGAA  4260
          AAAGATCAAG TTGTCATTTT AGAACAGAGG GAACTTGGGG AGAAAGCAGC CCAAGTAGGG  4320
          TATTTGTACA GTCAGAGGGC AACAGGAAGA TGCAGGCCTT CAAGGGCAAG GAGAGGCCAC  4380
          AAGGAATATG GGTGGGAGTA AAAGCAACAT CGTCTGCTTC ATACTTTTTC CTAGGCTTGG  4440
          CACTGCCTTT TCCTTTCTCA GGCCAATGGC AACTGCCATT TGAGTCCGGT GAGGGATCAG  4500
          CCAACCTCTT CTCTATGGCT CACCTTATTT GGAGTGAGAA ATCAAGGAGA CAGAGCTGAC  4560
          TGCATGATGA GTCTGAAGGC ATTTGCAGGA TGAACCTGAA CTGGTTGTGC AGAACAAACA  4620
          AGGCATTCAT GGGAATTGTT GTATTCCTTC TGCAGCCCTC CTTCTGGGCA CTAAGAAGGT  4680
          CTATGAATTA AATGCCTATC TAAAATTCTG ATTTATTCCT ACATTTTCTG TTTTTCTAATT  4740
          TGACCCTAAA ATCTATCGTGT TTTAGACTTA GACTTTTTAT TGCCCCCCCC CCCTTTTTTT  4800
          TTGAGACGGA GTCTCGCTCT GACGCACAGG CTGGAGTGCA GTGGCTCCGA TCTCTGCTCA  4860
          CTGAAAGCTC CGCCTCCCGG GTTCATGCCA TTCTCCTGCC TCAGCCTCCT GAGTAGCTGG  4920
          GACTACAGGC GCCCACCACC ACGCCCGGCT AATTTTTTGT ATTTTTAATA GAGACGGGGT  4980
```

```
TTCACTGTGT TAGCCAGGAT GGTCTCGATC TCCTGACCTC GTGATCCGCC TGCCTCGGCC   5040
TCCCAAAGTG CTGGGATTAC AGGCATGACC CACCGCTCCC GGCCTTGTTT TCCGTTTAAA   5100
GTCGTCTTCT TTTAATGTAA TCATTTTGAA CATGTGTGAA AGTTGATCAT ACGAATTGGA   5160
TCAATCTTGA AATACTCAAC CAAAAGACAG TCGAGAAGCC AGGGGGGAGAA AGAACTCAGG   5220
GCACAAAATA TTGGTCTGAG AATGGAATTC TCTGTAAGCA TAGTTGCTGA AATTTCCTGC   5280
TGTAACCAGA AGCCAGTTTT ATCTAACCGC TACTGAAACA CCCACTGTGT TTTGCTCACT   5340
CCCACTCACC GATCAAAACC TGCTACCTCC CCAAGACTTT ACTAGTGCCG ATAAACTTTC   5400
TCAAAGAGCA ACCAGTATCA CTTCCCTGTT TATAAAAACCT CTAACCATCT CTTTGTTCTT   5460
TGAACATGCT GAAAACCACC TGGTCTGCAT GTATGCCCGA ATTTGTAATT CTTTTCTCTC   5520
AAATGAAAAT TTAATTTTAG GGATTCATTT CTATATTTTC ACATATGTAG TATTATTATT   5580
TCCTTATATG TGTAAGGTGA AATTTATGGT ATTTGAGTGT GCAAGAAAAT ATATTTTTAA   5640
AGCTTTCATT TTTCCCCCAG TGAATGATTT AGAATTTTTT ATGTAAATAT ACAGAATGTT   5700
TTTTCTTACT TTTATAAGGA AGCAGCTGTC TAAAATGCAG TGGGGTTTGT TTTGCAATGT   5760
TTTAAACAGA GTTTTAGTAT TGCTATTAAA AGAAGTTACT TTGCTTTTAA AGAAACTTGG   5820
CTGCTTAAAA TAAGCAAAAA TTGGATGCAT AAAGTAATAT TTACAGATGT GGGGGAGATGT   5880
AATAAAACAA TATTAACTTG GCTGCTTAAA ATAAGCAAAA ATTGGATGCA TAAAGTAATA   5940
TTTACAGATG TGGGGAGATG TAATAAAACA ATATTAACTT GGTTTCTTGT TTTTGCTGTA   6000
TTTAGAGATT AAATAATTCT AAGATGATCA CTTTGCAAAA TTATGCTTAT GGCTGGCATG   6060
GAAATAGAAA TACTCAATTA TGTCTTTGTT GTATTAATGG GGAATATTTT GGACAATGTT   6120
TCATTATCAA ATTGTCGACA TCATTAATAT ATATTGTAAT GTTGGGAAGA GATCACTATT   6180
TTGAAGCACA GCTTTACAGA TCAGTATCTA TGATACATAT GTATAATAAA TTTTGATCGG   6240
GTATTAAAAG TATTAGAAGG TGGTTATAAT TGCAGAGTAT TCCATGAATA GTACACTGAC   6300
ACAGGGGTTT TACTTTGAGG ACCAGTGTAG TCAAGGGAAA ACATGAGTTA AAAAGAAAAG   6360
CAGGCAAATA TGCAGTCTTG ATTCTGCCAC TTACAGGATA GATAATGCCT GAACTTTAAT   6420
GACAAGATGA TCCAACCATA AAGGTGCTCT GTGCTTCACA GTGAATCTTT TCCCCATGCA   6480
GGAGTGTGCT CCCCTACAAA CGTTAAGACT GATCATTTCA AAAATCTATT AGCTATATCA   6540
AAAGCCTTAC ATTTTAATAT AGGTTGAACC AAAATTTCAA TTCCAGTAAC TTCTATTGTA   6600
ACCATTATTT TGTGTATGT CTTCAAGAAT GTTCATTGGA TTTTTGTTTG TAATAGTAAA   6660
ATACCGGATA CATTTCACGT GTCCTTCAGT ATTGATTTGG TTGAATATTG GGTCATAATG   6720
GTTGAGAAGC ATGGACACTA GAGCCAGAAT GCTTGGATAT GAATCCTGGA TCTGTCACTT   6780
ACTTCTGTGT GACCTTTGAA AGGCTACTTA TTTCCTCTCT ATTAAAATCA   6840
ATGAACAATG CCAGCCTCAT GGGGTTGTTG AATGATTAAA TTAGTTAATA TACCTAAAGT   6900
ACATAGAACA CTGCCTGCAC ATAGTAAAAG AATTATAAGT GTGAGGTAGT TGGTAAAATT   6960
ATGTAGTTGG ATATACTACC GAACAATATC TAATCTCTTT TTAGGGAAAT AAAGTTTGTG   7020
CATATATATA ATCCCGAAAC ATG
```

Seq ID NO: 326 Protein sequence
Protein Accession #: NP_077741.1

```
1           11          21          31          41          51
|           |           |           |           |           |
MAAAGPRRSV  RGAVCLHLLL  TLVIFSRDGS  ACKKVILNVP  SKLEADKIIG  RVNLEECFRS   60
ADLIRSSDPD  FRVLNDGSVY  TARAVALSDK  KRSFTIWLSD  KRKQTQKEVT  VLLEHQKVS   120
KTRHTRETVL  RRAKRRNAPI  PCSMQENSLG  PFPLFLQQVE  SDAAQNYTVF  YSISGRGVDK  180
EPLNLFYIER  DTGNLFCTRP  VDREEYDVFD  LIAYASTADG  YSADLPLPLP  IRVEDENDNH  240
PVFTEAIYNF  EVLESSRPGT  TVGVVCATDR  DEPDTMHTRL  KYSILQQTPR  SPGLFSVHPS  300
TGVITTVSRY  LDREVVDKYS  LIMKVQDMDG  QPFGLIGTST  CIITVTDSND  NAPTFRQNAY  360
EAFVEENAFN  VEILRIPIED  KDLINTANWR  VNFTILKGNE  NGHFKISTDK  ETNEGVLSVV  420
KPLNYEENRQ  VNLEIGVNNE  APFARDIPRV  TALNRALVTV  HVRDLDEGPE  CTPAAQYVRI  480
KENLAVGSKI  NGYKAYDPEN  RNGNGLRYKK  LEDPKGWITI  DEISGSIITS  KILDREVETP  540
KNELYNITVL  AIDKDDRSCT  GTLAVNIEDV  NDNPPEILQE  YVVICKPKMG  YTDILAVDPD  600
EPVHGAPPYF  SLPNTSPEIS  RLNSLTKVND  TAARLSYQRN  AGFQEYTIPI  TVKDRAGQAA  660
TKLLEVNLCE  CTHFTQCRAT  SRSTGVILGK  WAILAILLGI  ALLFSVLLTL  VCGVFGATKG  720
KRFPEDLAQQ  NLIISNTEAP  GDDRVCSANG  FMTQTTNNSS  QGFCGTMGSG  MKNGGQETIE  780
MMKGGNQTLE  SCRGAGHHHT  LDSCRGGHTE  VDNCRYTYSE  WHSFTQPRLG  EESIRGHTG
```

Seq ID NO: 327 DNA sequence
Nucleic Acid Accession #: NM_001941.2
Coding sequence: 64..2754

```
1           11          21          31          41          51
|           |           |           |           |           |
GGCAGGTCTC  GCTCTCGGCA  CCCTCCCGGC  GCCCGCGTTC  TCCTGGCCCT  GCCCGGCATC   60
CCGATGGCCG  CCGCTGGGCC  CCGGCGCTCC  GTGCGCGGAG  CGGTCTGCCT  GCATCTGCTG   120
CTGACCCTCG  TGATCTTCAG  TCGTGATGGT  GAAGCCTGCA  AAAAGGTGAT  ACTTAATGTA   180
CCTTCTAAAC  TAGAGGCAGA  CAAAATAATT  GGCAGAGTTA  ATTTGGAAGA  GTGCTTCAGG   240
TCTGCAGACC  TCATCCGGTC  AAGTGATCCT  GATTTCAGAG  TTCTAAATGA  TGGGTCAGTG   300
TACACAGCCA  GGGCTGTTGC  GCTGTCTGAT  AAGAAAAGAT  CATTTACCAT  ATGGCTTTCT   360
GACAAAAGGA  AACAGACACA  GAAAGAGGTT  ACTGTGCTGC  TAGAACATCA  GAAGGAGGTA   420
TCGAAGACAA  GACACACTAG  AGAAACTGTT  CTCAGGCGTG  CCAAGAGGAG  ATGGGCACCT   480
ATTCCTTGCT  CTATGCAAGA  GAATTCCTTG  GGCCCTTTCC  CATTGTTTCT  TCAACAAGTT   540
GAATCTGATG  CAGCACAGAA  CTATACTGTC  TTCTACTCAA  TAAGTGGACG  TGGAGTTGAT   600
AAAGAACCTT  TAAATTTGTT  TTATATAGAA  AGAGACACTG  GAAATCTATT  TTGCACTCGG   660
CCTGTGGATC  GTGAAGAATA  TGATGTTTTT  GATTTGATTG  CTTATGCGTC  AACTGCAGAT   720
GGATATTCAG  CAGATCTGCC  CCTCCCACTA  CCCATCCGGG  TAGAGGATGA  AAATGACAAC   780
CACCCTGTTT  TCACAGAAGC  AATTTATAAT  TTTGAAGTTT  TGGAAAGTAG  TAGACCTGGT   840
ACTACAGTGG  GGGTGGTTTG  TGCCACAGAC  AGAGATGAAC  CGGACACAAT  GCATACGCGC   900
CTGAAATACA  GCATTTTGCA  GCAGACACCA  AGGTCACCTG  GGCTCTTTTC  TGTGCATCCC   960
AGCACAGGCG  TAATCACCAC  AGTCTCTCAT  TATTTGGACA  GAGAGGTTGT  AGACAAGTAC  1020
TCATTGATAA  TGAAAGTACA  AGACATGGAT  GGCCAGTTTT  TTGGATTGAT  AGGCACATCA  1080
ACTTGTATCA  TAACAGTAAC  AGATTCAAAT  GATAATGCAC  CCACTTTCAG  ACAAAATGCT  1140
TATGAAGCAT  TTGTAGAGGA  AAATGCATTC  AATGTGGAAA  TCTTACGAAT  ACCTATAGAA  1200
GATAAGGATT  TAATTAACAC  TGCCAATTGG  AGAGTCAATT  TTACCATTTT  AAAGGGAAAT  1260
GAAAATGGAC  ATTTCAAAAT  CAGCACAGAC  AAAGAAACTA  ATGAAGGTGT  GTTTTCTGTT  1320
```

```
GTAAAGCCAC TGAATTATGA AGAAAACCGT CAAGTGAACC TGGAAATTGG AGTAAACAAT  1380
GAAGCGCCAT TTGCTAGAGA TATTCCCAGA GTGACAGCCT TGAACAGAGC CTTGGTTACA  1440
GTTCATGTGA GGGATCTGGA TGAGGGGCCT GAATGCACTC CTGCAGCCCA ATATGTGCGG  1500
ATTAAAGAAA ACTTAGCAGT GGGGTCAAAG ATCAACGGCT ATAAGGCATA TGACCCCGAA  1560
AATAGAAATG GCAATGGTTT AAGGTACAAA AAATTGCATG ATCCTAAAGG TTGGATCAAC  1620
ATTGATGAAA TTTCAGGGTC AATCATAACT TCCAAAATCC TGGATAGGGA GGTTGAAACT  1680
CCCAAAAATG AGTTGTATAA TATTACAGTC CTGGCAATAG ACAAAGATGA TAGATCATGT  1740
ACTGGAACAC TTGCTGTGAA CATTGAAGAT GTAAATGATA ATCCACCAGA AATACTTCAA  1800
GAATATGTAG TCATTTGCAA ACCAAAAATG GGGTATACCG ACATTTTAGC TGTTGATCCT  1860
GATGAACCTG TCCATGGAGC TCCATTTTAT TTCAGTTTGC CCAATACTTC TCCAGAAATC  1920
AGTAGACTGT GGAGCCTCAC CAAAGTTAAT GATACAGCTG CCCGTCTTTC ATATCAGAAA  1980
AATGCTGGAT TTCAAGAATA TACCATTCCT ATTACTGTAA AAGACAGGGC CGGCCAAGCT  2040
GCAACAAAAT TATTGAGAGT TAATCTGTGT GAATGTACTC ATCCAACTCA GTGTCGTGCG  2100
ACTTCAAGGA GTACAGGAGT AATACTTGGA AAATGGGCAA TCCTTGCAAT ATTACTGGGT  2160
ATAGCCACTG CTCTTTTCTGT ATTGCTAACT TTAGTATGTG GAGTTTTTGG TGCAACTAAA  2220
GGGAAACGTT TTCCTGAAGA TTTAGCACAC AAAACTTAA TTATATCAAA CACAGAAGCA  2280
CCTGGCGACG ATAGAGTGTG CTCTGCCAAT GGATTTATGA CCCAAACTAC CAACAACTCT  2340
AGCCAAGGTT TTTGTGGTAC TATGGGATCA GGAATGAAAA ATGGAGGGCA GGAAACCATT  2400
GAAATGATGA AAGGAGGAAA CCAGACCTTG GAATCCTGCC GGGGGGCTGG GCATCATCAT  2460
ACCCTGGACT CCTGCAGGGG AGGACACACG GAGGTGGACA ACTGCAGATA CACTTACTCG  2520
GAGTGGCACA GTTTTACTCA ACCCCGTCTC GGTGAAAAAT TGCATCGATG TAATCAGAAT  2580
GAAGACCGCA TGCCATCCCA AGATTATGTC CTCACTTATA ACTATGAGGG AAGAGGATCT  2640
CCAGCTGGTT CTGTGGGCTG CTGCAGTGAA AAGCAGGAAG AAGATGGCCT TGACTTTTTA  2700
AATAATTTGG AACCCAAATT TATTACATTA GCAGAAGCAT GCACAAAGAG ATAATGTCAC  2760
AGTGCTACAA TTAGGTCTTT GTCAGACATT CTGGAGGTTT CCAAAAATAA TATTGTAAAG  2820
TTCAATTTCA ACATGTATGT ATATGATGAT TTTTTTCTCA ATTTTGAATT ATGCTACTCA  2880
CCAATTTATA TTTTTAAAGC CAGTTGTTGC TTATCTTTTC CAAAAAGTGA AAAATGTTAA  2940
AACAGACAAC TGGTACATCT CAAACTCCAG CACTGGAATT AAGGTCTCTA AAGCATCTGC  3000
TCTTTTTTTT TTTTACGGAT ATTTTAGTAA TAAATATGCT GGATAAATAT TAGTCCAACA  3060
ATAGCTAAGT TATGCTAATA TCACATTATT ATGTATTCAC TTTAAGTGAT AGTTTAAAAA  3120
ATAAACAAGA AATATTGAGT ATCACTATGT GAAGAAAGTT TTGGAAAAGA AACAATGAAG  3180
ACTGAATTAA ATTAAAAATG TTGCAGCTCA TAAAGAATTG GGACTCACCC CTACTGCACT  3240
ACCAAATTCA TTTGACTTTG GAGGCAAAAT GTGTTGAAGT GCCCTATGAA GTAGCAATTT  3300
TCTATAGGAA TATAGTTGGA AATAAATGTG TGTGTGTATA TTATTATTAA TCAATGCAAT  3360
ATTTAAAATG AAATGAGAAC AAAGAGGAAA ATGGTAAAAA CTTGAAATGA GGCTGGGGTA  3420
TAGTTTGTCC TACAATAGAA AAAAGAGAGA GCTTCCTAGG CCTGGGCTCT TAAATGCTGC  3480
ATTATAACTG AGTCTATGAG GAAATAGTTC CTGTCCAATT TGTGTAATTT GTTTAAAATT  3540
GTAAATAAAT TAAACTTTTC TGGTTTCTGT GGGAAGGAAA TAGGGAATCC AATGGAACAG  3600
TAGCTTTGCT TTGCAGTCTG TTTCAAGATT TCTGCATCCA CAAGTTAGTA GCAAACTGGG  3660
GAATACTCGC TGCAGCTGGG GTTCCCTGCT TTTTGGTAGC AAGGGTCCAG AGATGAGGTG  3720
TTTTTTTCGG GGAGCTAATA ACAAAAACAT TTTAAAACTT ACCTTTACTG AAGTTAAATC  3780
CTCTATTGCT GTTTCTAATC TCTCTTATAG TGACCAACAT CTTTTTAATT TAGATCCAAA  3840
TAACCATGTC CTCCTAGAGT TTAGAGGCTA GAGGGAGCTG AGGGGAGGAT CTTACTGAAA  3900
GCACCCTGGG GAGATTGATT GTCCTTAAAC CTAAGCCCCA CAAACTTGAC ACCTGATCAG  3960
GTCTGGGAGC TACAAAATTT CATTTTTCTC CTCACTGCCC TTCTTCTGAG TGGCATTGGC  4020
CTGAATCAAG GAAAGCCAGG CCTTGTGGGC CCCCTTCTTT CGGCTTTCTG CTAAAGCAAC  4080
ACCTCCAGCA GAGATTCCCT TAAGTGACTC CAGGTTTTCC ACCATCCTTC AGCGTGAATT  4140
AATTTTTAAT CAGTTTGCTT TCTCCAGAGA AATTTTTAAAA TAATAGAAGA AATAGAAATT  4200
TTGAATGTAT AAAAGAAAAA GATCAAGTTG TCATTTTAGA ACAGAGGGAA CTTTGGGAGA  4260
AAGCAGCCCA AGTAGGTTAT TTGTACAGTC AGAGGGCAAC AGGAAGATGC AGGCCTTCAA  4320
GGGCAAGGAG AGGCCACAAG GAATATGGGT GGGAGTAAAA GCAACATCGT CTGCCTTCATA  4380
CTTTTTCCTA GGCTTGGCAC TGCCCTTTCC TTTCTCAGGC CAATGGCAAC TGCCATTTGA  4440
GTCCGGTGAG GGATCAGCCA ACCTCTTCTC TATGGCTCAC CTTATTTGGA GTGAGAAATC  4500
AAGGAGACAG AGCTGACTGC ATGATGAGTC TGAAGGCATT TGCAGGATGA GCCTGAACTG  4560
GTTGTGCAGA ACAAACAAGG CATTCATGGG AATTGTTGTA TTCCTTCTGC AGCCCTCCTT  4620
CTGGGCACTA AGAAGGTCTA TGAATTAAAT GCCTATCTAA AATTCTGATT TATTCCTACA  4680
TTTTCTGTTT TCTAATTTGA CCCTAAAATC TATGTGTTTT AGACTTAGAC TTTTTATTGC  4740
CCCCCCCCCC TTTTTTTTTG AGACGGAGTC TCGCTCTGAC GCACAGGCTG GAGTGCAGTG  4800
GCTCCGATCT CTGCTCACTG AAAGCTCCGC CTCCCGGGTT CATGCCATTC TCCTGCCTCA  4860
GCCTCCTGAG TAGCTGGGAC TACAGGCGCC CACCACCACG CCCGGCTAAT TTTTTGTATT  4920
TTTAATAGAG ACGGGGTTTC ACTGTGTTAG CCAGGATGGT CTCGATCTCC TGACCTCGTG  4980
ATCCGCCTGC CTCGGCCTCC CAAAGTGCTG GGATTACAGG CATGACCCAC CGCTCCCGGC  5040
CTTGTTTTCC GTTAAAGTC GTCTTCTTTT AATGTAATCA TTTTGAACAT GTGTGAAAGT  5100
TGATCATACG AATTGGATCA ATCTTGAAAT ACTCAACCAA AAGACAGTCG AGAAGCCAGG  5160
GGGAGAAAGA ACTCAGGGCA CAAAATATTG GTCTGAGAAT GGAATTCTCT GTAAGCCTAG  5220
TTGCTGAAAT TTCCTGCTGT AACCAGAAGC CAGTTTTATC TAAACGGCTAC TGAAACACCC  5280
ACTGTGTTTT GCTCACTCCC TCACTCACCG ATCAAAACCT GCTACCTCCC CAAGACTTTA  5340
CTAGTGCCGA TAAACTTTCT CAAAGAGCAA CCAGTATCAC TTCCCTGTTT ATAAAACCTC  5400
TAACCATCTC TTTGTTCTTT GAACATGCTG AAAACCACCT GGTCTGCATG TATGCCCTAA  5460
TTTGTAATTC TTTTCTCTCA AATGAAAATT TAATTTTAGG GATTCATTTC TATATTTTCA  5520
CATATGTAGT ATTATTATTT CCTTATATGT GTAAGGTGAA ATTTATGGTA TTTGAGTGTG  5580
CAAGAAAATA TATTTTTAAA GCTTTCATTT TTCCCCCAGT GAATGATTTA GAATTTTTTA  5640
TGTAAATATA CAGAATGTTT TTTCTTACTT TTATAAGGAA GCAGCTGTCT AAAATGCAGT  5700
GGGGTTTGTT TTGCAATGTT TTAAACAGAG TTTTAGTATT GCTATTAAAA GAAGTTACTT  5760
TGCTTTTAAA GAAACTTGGC TGCTTAAAAT AAGCAAAAAT TGGATGCATA AAGTAATATT  5820
TACAGATGTG GGGAGATGTA ATAAAACAAT ATTAACTTGG TTTCTTGTTT TTGCTGTATT  5880
TAGAGATTAA ATAATTCTAA GATGATCACT TTGCAAAATT ATGCTTATGG CTGGCATGGA  5940
AATAGAAATA CTCAATTATG TCTTTGTTGT ATTAATGGGG AATATTTTGG ACAATGTTTC  6000
ATTATCAAAT TGTCGACATC ATTAATATAT ATTGTAATGT TGGGAAGAGA TCACTATTTT  6060
GAAGCACAGC TTTACAGATG AGTATCTATG ATACATATGT ATAATAAATT TTGATCGGGT  6120
ATTAAAAGTA TTAGAAGGTG GTTATAATTG CAGATATTC CATGAATAGT ACACTGACAC  6180
AGGGGTTTTA CTTTGAGGAC CAGTGTAGTC AAGGGAAAAC ATGAGTTAAA AAGAAAAGCA  6240
GGCAATATTG CAGTCTTGAT TCTGCCACTT ACAGGATAGA TAATGCCTGA ACTTTAATGA  6300
CAAGATGATC CAACCATAAA GGTGCTCTGT GCTTCACAGT GAATCTTTTC CCCATGCAGG  6360
```

```
AGTGTGCTCC CCTACAAACG TTAAGACTGA TCATTTCAAA AATCTATTAG CTATATCAAA  6420
AGCCTTACAT TTTAATATAG GTTGAACCAA AATTTCAATT CCAGTAACTT CTATTGTAAC  6480
CATTATTTTT GTGTATGTCT TCAAGAATGT TCATTGGATT TTTGTTTGTA ATAGTAAAAT  6540
ACCGGATACA TTTCACGTGT CCTTCAGTAT TGATTTGGTT GAATATTGGG TCATAATGGT  6600
TGAGAAGCAT GGACACTAGA GCCAGAATGC TTGGATATGA ATCCTGGATC TGTCACTTAC  6660
TTCTGTGTGA CCTTTGAAAG GCTACTTATT TCCTCTCTTA GCTTTCTCAT TAAAATCAAT  6720
GAACAATGCC AGCCTCATGG GGTTGTTGAA TGATTAAATT AGTTAATATA CCTAAAGTAC  6780
ATAGAACACT GCCTGCACAT AGTAAAAGAA TTATAAGTGT GAGGTAGTTG GTAAAATTAT  6840
GTAGTTGGAT ATACTACCGA ACAATATCTA ATCTCTTTTT AGGGAAATAA AGTTTGTGCA  6900
TATATATAAT CCCGAAACAT G
```

Seq ID NO: 328 <u>Protein sequence</u>
Protein Accession #: NP_001932.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MAAAGPRRSV RGAVCLHLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECFRS   60
ADLIRSSDPD FRVLNDGSVY TARAVALSDK KRSFTIWLSD KRKQTQKEVT VLLEHQKKVS  120
KIRHTRETVL RRAKRRWAPI PCSMQENSLG PFPLFLQQVE SDAAQNYTVP YSISGRGVDK  180
EPLNLFYIER DTGNLFCTRP VDREEYDVFD LIAYASTADG YSADLPLPLP IRVEDENDNH  240
PVFTEAIYNF EVLESSRPGT TVGVVCATDR DEPDTMHTRL KYSILQQTPR SPGLFSVHPS  300
TGVITTVSHY LDREVVDKYS LIMKVQDMDG QFFGLIGTST CIITVTDSND NAPTFRQNAY  360
EAFVEENAFN VEILRIPIED KDLINTANWR VNFTILKGNE NGHPKISTDK ETNEGVLSVV  420
KPLNYEENRQ VNLEIGVNNE APPARDIPRV TALNRALVTV HVRDLDEGPE CTPAAQYVRI  480
KENLAVGSKI NGYKAYDPEN RNGNGLRYKK LHDPKGWITI DEISGSIITS KILDREVETP  540
KNELYNITVL AIDKDDRSCT GTLAVNIEDV NDNPPEILQE YVVICKPKMG YTDILAVDPD  600
EPVHGAPFYF SLFNTSPEIS RLWSLTKVND TAARLSYQKN AGFQEYTIPI TVKDRAGQAA  660
TKLLRVNLCE CTHPTQCRAT SRSTGVILGK WAILAILLGI ALLFSVLLTL VCGVFGATKG  720
KRFPEDLAQQ NLIISNTEAP GDDRVCSANG FMTQTTNNSS QGFCGTMGSG MKNGGQETIE  780
MMKGGNQTLE SCRGAGHHHT LDSCRGGHTE VDNCRYTYSE WHSPTQPRLG EKLHRCNQNE  840
DRMPSQDYVL TYNYEGRGSP AGSVGCCSEK QEEDGLDFLN NLEPKFITLA EACTKR
```

Seq ID NO: 329 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_016583.2
Coding sequence: 72..842

```
1          11         21         31         41         51
|          |          |          |          |          |
GGAGTGGGGG AGAGAGAGGA GACCAGGACA GCTGCTGAGA CCTCTAAGAA GTCCAGATAC   60
TAAGAGCAAA GATGTTTCAA ACTGGGGGCC TCATTGTCTT CTACGGGCTG TTAGCCCAGA  120
CCATGGCCCA GTTTGGAGGC CTGCCCGTGC CCCTGGACCA GACCCTGCCC TTGAATGTGA  180
ATCCAGCCCT GCCCTTGAGT CCCACAGGTC TTGCAGGAAG CTTGACAAAT GCCCTCAGCA  240
ATGGCCTGCT GTCTGGGGGC CTGTTGGGCA TTCTGGAAAA CCTTCCGCTC CTGGACATCC  300
TGAAGCCTGG AGGGAGGTACT TCTGGTGGCC TCCTTGGGGG ACTGCTTGGA AAAGTGACGT  360
CAGTGATTCC TGGCCTGAAC AACATCATTG ACATAAAGGT CACTGACCCC CAGCTGCTGG  420
AACTTGGCCT TGTGCAGAGC CCTGATGGCC ACCGTCTCTA TGTCACCATC CCTCTCGGCA  480
TAAAGCTCCA AGTGAATACG CCCCTGGTCG GTGCAAGTCT GTTGAGGCTG GCTGTGAAGC  540
TGGACATCAC TGCAGAAATC TTAGCTGTGA GAGATAAGCA GGAGAGGATC CACCTGGTCC  600
TTGGTGACTG CACCCATTCC CCTGGAAGCC TGCAAATTTC TCTGCTTGAT GGACTTGGCC  660
CCCTCCCCAT TCAAGGTCTT CTGGACAGCC TCACAGGGAT CTTGAATAAA GTCCTGCCTG  720
AGTTGGTTCA GGGCAACGTG TGCCCTCTGG TCAATGAGGT TCTCAGAGGC TTGGACATCA  780
CCCTGGTGCA TGACATTGTT AACATGCTGA TCCACGGACT ACAGTTTGTC ATCAAGGTCT  840
AAGCCTTCCA GGAAGGGGCT GGCCTCTGCT GAGCTGCTTC CCAGTGCTCA CAGATGGCTG  900
GCCCATGTGC TGGAAGATGA CACAGTTGCC TTCTCTCCGA GGAACCTGCC CCCTCTCCTT  960
TCCCACCAGG CGTGTGTAAC ATCCCATGTG CCTCACCTAA TAAAATGGCT CTTCTTCTGC 1020
AAAAAAAAAA AAAAAAAAAA AAAAAAAAA
```

Seq ID NO: 330 <u>Protein sequence</u>
Protein Accession #: NP_057667.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MFQTGGLIVF YGLLAQTMAQ FGGLPVPLDQ TLPLNVNPAL PLSPTGLAGS LTNALSNGLL   60
SGGLLGILEN LPLLDILKPG GGTSGGLLGG LLGKVTSVIP GLNNIIDIKV TDPQLLELGL  120
VQSPDGHRLY VTIPLGIKLQ VNTPLVGASL LRLAVKLDIT AEILAVRDKQ ERIHLVLGDC  180
THSPGSLQIS LLDGLGPLPI QGLLDSLTGI LNKVLPELVQ GNVCPLVNEV LRGLDITLVH  240
DIVNMLIHGL QFVIKV
```

Seq ID NO: 331 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_004363.1
Coding sequence: 115..2223

```
1          11         21         31         41         51
|          |          |          |          |          |
CTCAGGGCAG AGGGAGGAAG GACAGCAGAC CAGACAGTCA CAGCAGCCTT GACAAAACGT   60
TCCTGGAACT CAAGCTCTTC TCCACAGAGG AGGACAGAGC AGACAGCAGA GACCATGGAG  120
TCTCCCTGG CCCCTCCCCA CAGATGGTGC ATCCCCTGGC AGAGGCTCCT GCTCACAGCC  180
TCACTTCTAA CCTTCTGGAA CCCGCCCACC ACTGCCAAGC TCACTATTGA TCCACGCCG  240
TTCAATGTCG CAGAGGGGAA GGAGGTGCTT CTACTTGTCC ACAATCTGCC CAGCATCTT  300
TTTGGCTACA GCTGGTACAA AGGTGAAAGA GTGGATGGCA ACCGTCAAAT TATAGGGATT  360
GTAATAGGAA CTCAACAAGC TACCCCAGGG CCCGCATACA GTGGTCGAGA GATAATATAC  420
CCCAATGCAT CCCTGCTGAT CCAGAACATC ATCCAGAATG ACACAGGATT CTACACCCTA  480
```

```
CACGTCATAA AGTCAGATCT TGTGAATGAA GAAGCAACTG GCCAGTTCCG GGTATACCCG    540
GAGCTGCCCA AGCCCTCCAT CTCCAGCAAC AACTCCAAAC CGTGGAGGA CAAGGATGCT    600
GTGGCCTTCA CCTGTGAACC TGAGACTCAG GACGCAACCT ACCTGTGGTG GGTAAACAAT    660
CAGAGCCTCC CGGTCAGTCC CAGGCTGCAG CTGTCCAATG GCAACAGGAC CCTCACTCTA    720
TTCAATGTCA CAAGAAATGA CACAGCAAGC TACAAATGTG AAACCCAGAA CCCAGTGAGT    780
GCCAGGCGCA GTGATTCAGT CATCCTGAAT GTCCTCTATG GCCCGGATGC CCCCACCATT    840
TCCCCTCTAA ACACATCTTA CAGATCAGGG GAAAATCTGA ACCTCTCCTG CCACGCAGCC    900
TCTAACCCAC CTGCACAGTA CTCTTGGTTT GTCAATGGGA CTTTCCAACA ATCCACCCAA    960
GAGCTCTTTA TCCCCAACAT CACTGTGAAT AATAGTGGAT CCTATACGTG CCAAGCCCAT   1020
AACTCAGACA CTGGCCTCAA TAGGACCACA GTCACGACGA TCACAGTCTA TGCAGAGCCA   1080
CCCAAACCCT TCATCACCAG CAACAACTCC AACCCCGTGG AGGATGAGGA TGCTGTAGCC   1140
TTAACCTGTG AACCTGAGAT TCAGAACACA ACCTACCTGT GGTGGGTAAA TAATCAGAGC   1200
CTCCCGGTCA GTCCCAGGCT GCAGCTGTCC AATGACAACA GGACCCTCAC TCTACTCAGT   1260
GTCACAAGGA ATGATGTAGG ACCCTATGAG TGTGGAATCC AGAACGAATT AAGTGTTGAC   1320
CACAGCGACC CAGTCATCCT GAATGTCCTC TATGGCCCAG ACGACCCCAC CATTTCCCCC   1380
TCATACACCT ATTACCGTCA AGGGGTGAAC CTCAGCCTCT CCTGCCATGC AGCCTCTAAC   1440
CCACCTGCAC AGTATTCTTG GCTGATTGAT GGGAACATCC AGCAACACAC ACAAGAGCTC   1500
TTTATCTCCA ACATCACTGA GAAGAACAGC GGACTCTATA CCTGCCAGGC CAATAACTCA   1560
GCCAGTGGCC ACAGCAGGAC TACAGTCAAG ACAATCACAG TCTCTGCGGA GCTGCCCAAG   1620
CCCTCCATCT CCAGCAACAA CTCCAAACCC GTGGAGGACA AGGATGCTGT GGCCTTCACC   1680
TGTGAACCTG AGGCTCAGAA CACAACCTAC CTGTGGTGGG TAAATGGTCA GAGCCTCCCA   1740
GTCAGTCCCA GGCTGCAGCT GTCCAATGGC AACAGGACCC TCACTCTATT CAATGTCACA   1800
AGAAATGACG CAAGAGCCTA TGTATGTGGA ATCCAGAACT CAGTGAGTGC AAACCGCAGT   1860
GACCCAGTCA CCCTGGATGT CCTCTATGGG CCGGACACCC CCATCATTTC CCCCCCAGAC   1920
TCGTCTTACC TTTCGGGAGC GAACCTCAAC CTCTCCTGCC ACTCGGCCTC TAACCCATCC   1980
CCGCAGTATT CTTGGCGTAT CAATGGGATA CCGCAGCAAC ACACACAAGT TCTCTTTATC   2040
GCCAAAATCA CGCCAAATAA TAAGGGGACC TATGCCTGTT TTGTCTCTAA CTTGGCTACT   2100
GGCCGCAATA ATTCCATAGT CAAGAGCATC ACAGTCTCTG CATCTGGAAC TTCTCCTGGT   2160
CTCTCAGCTG GGGCCACTGT CGGCATCATG ATTGGAGTGC TGGTTGGGGT TGCTCTGATA   2220
TAGCAGCCCT GGTGTAGTTT CTTCATTTCA GGAAGACTGA CAGTTGTTTT GCTTCTTCCT   2280
TAAAGCATTT GCAACAGCTA CAGTCTAAAA TTGCTTCTTT ACCAAGGATA TTTACAGAAA   2340
AGACTCTGAC CAGAGATCGA GACCATCCTA GCCAACATCG TGAAACCCCA TCTCTACTAA   2400
AAATACAAAA ATGAGCTGGG CTTGGTGGCG CGCACCTGTA GTCCCAGTTA CTCGGGAGGC   2460
TGAGGCAGGA GAATCGCTTG AACCCGGGAG GTGGAGGATTG CAGTGAGCCC AGATCGCACC   2520
ACTGCACTCC AGTCTGGCAA CAGAGCAAGA CTCCATCTCA AAAGAAAAG AAAAGAAGAC   2580
TCTGACCTGT ACTCTTGAAT ACAAGTTTCT GATACCACTG CACTGTCTGA GAATTCCAA   2640
AACTTTAATG AAACTAACTGA CAGCTTCATG AAACTGTCCA CCAAGATCAA GCAGAGAAAA   2700
TAATTAATTT CATGGGACTA AATGAACTAA TGAGGATTGC TGATTCTTTA AATGTCTTGT   2760
TTCCCAGATT TCAGGAAACT TTTTTTCTTT TAAGCTATCC ACTCTTACAG CAATTTGATA   2820
AAATATACTT TTGTGAACAA AAATTGAGAC ATTTACATTT TCTCCCTATG TGGTGGCTCC   2880
AGACTTGGGA AACTATTCAT GAATATTTAT ATTGTATGGT AAATATAGTTA TTGCACAAGT   2940
TCAATAAAAA TCTGCTCTTT GTATAACAGA AAAA
```

Seq ID NO: 332 Protein sequence
Protein Accession #: NP_004354.1

```
1        11        21        31        41        51
|         |         |         |         |         |
MESPSAFPHR WCIPWQRLLL TASLLTFWNP PTTAKLTIES TPFNVAEGKE VLLLVHNLPQ    60
HLFGYSWYKG ERVDGNRQII GYVIGTQQAT PGPAYSGREI IYPNASLLIQ NIIQNDTGFY   120
TLHVIKSDLV NEEATGQFRV YPELPKPSIS SNNSKPVEDK DAVAFTCEPE TQDATYLWWV   180
NNQSLPVSPR LQLSNGNRTL TLFNVTRNDT ASYKCETQNP VSARRSDSVI LNVLYGPDAP   240
TISPLNTSYR SGENLNLSCH AASNPPAQYS WFVNGTFQQS TQELFIPNIT VNNSGSYTCQ   300
AHNSDTGLNR TTVTTITVYA EPPKPFITSN NSNPVEDEDA VALTCEPEIQ NTTYLWWVNN   360
QSLPVSPRLQ LSNDNRTLTL LSVTRNDVGP YECGIQNELS VDHSDPVILN VLYGPDDPTI   420
SPSYTYYRPG VNLSLSCHAA SNPPAQYSWL IDGNIQQETQ ELFISNITEK NSGLYTCQAN   480
NSASGHSRTT VKTITVSAEL PKPSISSNNS KPVEDKDAVA FTCEPEAQNT TYLWWVNGQS   540
LPVSPRLQLS NGNRTLTLFN VTRNDARAYV CGIQNSVSAN RSDPVTLDVL YGPDTPIISP   600
PDSSYLSGAN LNLSCHSASN PSPQYSWRIN GIPQQHTQVL FIAKITPNNN GTYACFVSNL   660
ATGRNNSIVK SITVSASGTS PGLSAGATVG IMIGVLVGVA LI
```

Seq ID NO: 333 DNA sequence
Nucleic Acid Accession #: NM_006952.1
Coding sequence: 11..793

```
1        11        21        31        41        51
|         |         |         |         |         |
AATCCCGACA ATGGCGAAAG ACAACTCAAC TGTTCGTTGC TTCCAGGGCC TGCTGATTTT    60
TGGAAATGTG ATTATTGGTT GTTGCGGCAT TGCCCTGACT GCGGAGTGCA TCTTCTTTGT   120
ATCTGACCAA CACAGCCTCT ACCCACTGCT TGAAGCCACC GACAACGATG ACATCTATGG   180
GGCTGCCTGG ATCGGCATAT TTGTGGGCAT CTGCCTCTTC TGCCTGTCTG TTCTAGGCAT   240
TGTAGGCATC ATGAAGTCCA GCAGGAAAAT TCTTCTGGCG TATTTCATTC TGATGTTTAT   300
AGTATATGCC TTTGAAGTGG CATCTTGTAT CACAGCAGCA ACACAACGAG ACTTTTTCAC   360
ACCCAACCTC TTCCTGAAGC AGATGCTAGA GAGGTACCAA AACAACAGCC CTCCAAACAA   420
TGATGACCAG TGGAAAAACA ATGGAGTCAC CAAAACCTGG GACAGGCTCA TGCTCCAGGA   480
CAATTGCTGT GGCGTAAATG GTCCATCAGA CTGGCAAAAA TACACATCTG CCTTCCGGAC   540
TGAGAATAAT GATGCTGACT ATCCCTGGCC TCGTCAATGC TGTGTTATGA ACAATCTTAA   600
AGAACCTCTC AACCTGGAGG CTTGTAAACT AGGCGTGCCT GGTTTTTATC ACAATCAGGG   660
CTGCTATGAA CTGATCTCTG GTCCAATGAA CCGACACGCC TGGGGGGTTG CCTGGTTTGG   720
ATTTGCCATT CTCTGCTGGA CTTTTTGGGT TCTCCTGGGT ACCATGTTCT ACTGGAGCAG   780
AATTGAATAT TAAGAA
```

Seq ID NO: 334 Protein sequence
Protein Accession #: NP_008883.1

```
       1         11        21        31        41        51
       |         |         |         |         |         |
5      MAKDWSTVRC FQGLLIFGNV IIGCCGIALT AECIFFVSDQ HSLYPLLEAT DNDDIYGAAW   60
       IGIFVGICLF CLSVLGIVGI MKSSRKILLA YPILMFIVYA PEVASCITAA TQRDFFTPNL  120
       FLKQMLERYQ NNSPPNNDDQ WKNNGVTKIW DRLMLQDNCC GVNGPSDWQK YTSAFRTENN  180
       DADYPWPRQC CVMNNLKEPL NLEACKLGVP GPYHNQGCYE LISGPMNRHA WGVAWFGFAI  240
       LCWTPWVLLG TMFYWSRIEY
```

10   Seq ID NO: 335 <u>DNA sequence</u>
      Nucleic Acid Accession #: NM_002638.1
      Coding sequence: 120..473

```
       1         11        21        31        41        51
       |         |         |         |         |         |
15     CAATACAGCT AAGGAATTAT CCCTTGTAAA TACCACAGAC CCGCCCTGGA GCCAGGCCAA   60
       GCTGGACTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA  120
       TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCCTCAT CGGCTGGGACG CTGGTTCTAG  180
       AGGCAGCTGT CACGGGAGTT CCTGTTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCCAT  240
20     TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAGTTAA AGGTCAAGAT AAAGTCAAAG  300
       CGCAAGAGCC AGTCAAAGGT CCAGTCTCCA CTAAGCCTGG CTCCTGCCCC ATTATCTTGA  360
       TCCGGTGCGC CATGTTGAAT CCCCCTAACC GCTGCTTGAA AGATACTGAC TGCCCAGGAA  420
       TCAAGAAGTG CTGTGAAGGC TCTTGCGGGA TGGCCTGTTT CGTTCCCCAG TGAAGGGAGC  480
       CGGTCCTTGC TGCACCTGTG CCGTCCCCAG AGCTACAGGC CCCATCTGGT CCTAAGTCCC  540
25     TGCTGCCCTT CCCCTTCCCA CACTGTCCAT TCTTCCTCCC ATTCAGGATG CCCACGGCTG  600
       GAGCTGCCTC TCTCATCCAC TTTCCAATAA A
```

      Seq ID NO: 336 <u>Protein sequence</u>
      Protein Accession #: NP_002629.1

30
```
       1         11        21        31        41        51
       |         |         |         |         |         |
       MRASSFLIVV VFLIAGTLVL EAAVTGVPVK GQDTVKGRVP FNGQDPVKGQ VSVKGQDKVK   60
       AQEPVKGPVS TKPGSCPIIL IRCAMLNPPN RCLKDTDCPG IKKCCEGSCG MACFVPQ
```

35   Seq ID NO: 337 <u>DNA sequence</u>
      Nucleic Acid Accession #: NM_001793.2
      Coding sequence: 71..2560

```
       1         11        21        31        41        51
40     |         |         |         |         |         |
       AAAGGGGCAA GAGCTGAGCG GAACACCGGC CCGCCGTCGC GGCAGCTGCT TCACCCCTCT   60
       CTCTGCAGCC ATGGGGCTCC CTCGTGGACC TCTCGCGTCT CTCCTCCTTC TCCAGGTTTG  120
       CTGGCTGCAG TGCGCGGCCT CCGGAGCCGTG CCGGGCGGTC TTCAGGGAGG CTGAAGTGAC  180
45     CTTGGAGGCG GGAGGCGCGG AGCAGGAGCC CGGCCAGGCG CTGGGGAAAG TATTCATGGG  240
       CTGCCCTGGG CAAGAGCCAG CTCTGTTTAG CACTGATAAT GATGACTTCA CTGTGCGGAA  300
       TGGCGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAAGG AATCCATTGA AGATCTTCCC  360
       ATCCAAACGT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCCAA TATCTGTCCC  420
       TGAAAATGGC AAGGGTCCCT TCCCCCAGAG ACTGAATCAG CTCAAGTCTA ATAAAGATAG  480
50     AGACACCAAG ATTTTCTACA GCATCACGGG GCCGGGGGCG GACAGCCCCC CTGAGGGTGT  540
       CTTCGCTGTA GAGAAGGGAG CAGGCTGGTT GTTGTTGAAT AAGCCACTGG ACCGGGAGGA  600
       GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTCAGAG AATGGTGCCT CAGTGGAGGA  660
       CCCCATGAAC ATCTCCATCA TCGTGACCGA CCAGAATGAC CACAAGCCCA AGTTTACCCA  720
       GGACACCTTC CGAGGGAGTG TCTTAGAGGG AGTCCTACCA GGTACTTCTG TGATGCAGGT  780
55     GACAGCCACG GATGAGGATG ATGCCATCTA CACCTACAAT GGGGTGGTTG CTTACTCCAT  840
       CCATAGCCAG GAACCAAAGG ACCCACACGA CCTCATGTTC ACCATTCACC GGAGCACAGG  900
       CACCATCAGC GTCATCTCCA GTGGCCTGGA CCGGGAAAAA GTCCCTGAGT ACACACTGAC  960
       CATCCAGGCC ACAGACATGG ATGGGGACGG CTCCACCACC ACGGCAGTGG CAGTAGTGGA 1020
       GATCCTTGAT GCCAATGACA ATGCTCCCAT GTTTGACCCC CAGAAGTACG AGGCCCATGT 1080
60     GCCTGAGAAT GCAGTGGGCC ATGAGGTGCA GAGGCTGACG GTCACTGATC TGGACGCCCC 1140
       CAACTCACCA GCGTGGCGTG CCACCTACCT TATCATGGGC GGTGACGACG GGGACCATTT .1200
       TACCATCACC ACCCACCCTG AGAGCAACCA GGGCATCCTG ACAACCAGGA AGGGTTTGGA 1260
       TTTTGAGGCC AAAAACCAGC ACACCCTGTA CGTTGAAGTG ACCAACGAGG CCCCTTTTGT 1320
       GCTGAAGCTC CCAACCTCCA CAGCCACCAT AGTGGTCCAC GTGGAGGATG TGAATGAGGC 1380
65     ACCTGTGTTT GTCCCACCCT CCAAAGTCGT TGAGGTCCAA GAGGGCATCC CCACTGGGGA 1440
       GCCTGTGTGT GTCTACACTG CAGAAGACCC TGACAAGGAG AATCAAAAGA TCAGCTACCG 1500
       CATCCTGAGA GACCCAGCAG GGTGGCTAGC CATGGACCCA GACAGTGGGC AGGTCACAGC 1560
       TGTGGGCACC CTCGACCGTG AGGATGAGCA GTTTGTGAGG AACAACATCT ATGAAGTCAT 1620
       GGTCTTGGCC ATGGACAATG GAAGCCCTCC CACCACTGGC ACGGGAACCC TTCTGCTAAC 1680
70     ACTGATTGAT GTCAATGACC ATGGCCCAGT CCCTGAGCCC CGTCAGATCA CCATCTGCAA 1740
       CCAAAGCCCT GTGCGCCAGG TGCTGAACAT CACGGACAAG GACCTGTCTC CCCACACCTC 1800
       CCCTTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGGACGGCAG AGGTCAACGA 1860
       GGAAGGTGAC ACAGTGGTCT TGTCCCTGAA GAAGTTCCTG AAGCAGGATA CATATGACGT 1920
       GCACCTTTCT CTGTCTGACC ATGGCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980
75     GTGCGACTGC CATGGCCATG TCGAAACCTG CCCTGGACCC TGGAAGGGAG GTTTCATCCT 2040
       CCCTGTGCTG GGGGCTGTCC TGGCTCTGCT GTTCCTCCTG CTGGTGCTGC TTTTGTTGGT 2100
       GAGAAAGAAG CGGAAGATCA AGGAGCCCCT CCTACTCCCA GAAGATGACA CCCGTGACAA 2160
       CGTCTTCTAC TATGGGGAAG AGGGGGGTGG CGAAGAGGAC CGAAGACTATG ACATCACCCA 2220
       GCTCCACCGA GGTCTGGAGG CCAGGCCGGA GGTGGTTCTC CGCAATGACG TGGCACCAAC 2280
80     CATCATCCCG ACACCCATGT ACCGTCCTCG GCCAGCCAAC CCAGATGAAA TCGGCAACTT 2340
       TATAATTGAG AACCTGAAGG CGGCTAACAC AGACCCCACA GCCCGGCCCT ACGACACCCT 2400
       CTTGGTTGTTC GACTATGAGG GCAGCGGCTC CGACGCCGCG TCCCTGAGCT CCCTCACCTC 2460
       CTCCGCCTCC GACCAAGACC AAGATTACGA TTATCTGAAC GAGTGGGGCA GCCGCTTCAA 2520
       GAAGCTGGCA GACATGTACG GTGGCGGGGA GGACGACTAG GCGGCCTGCC TGCAGGGCTG 2580
```

1083

```
      GGGACCAAAC GTCAGGCCAC AGAGCATCTC CAAGGGGTCT CAGTTCCCCC TTCAGCTGAG  2640
      GACTTCGGAG CTTGTCAGGA AGTGGCCGTA GCAACTTGGC GGAGACAGGC TATGAGTCTG  2700
      ACGTTAGAGT GGTTGCTTCC TTAGCCTTTC AGGATGGAGG AATGTGGGCA GTTTGACTTC  2760
   5  AGCACTGAAA ACCTCTCCAC CTGGGCCAGG GTTGCCTCAG AGGCCAAGTT TCCAGAAGCC  2820
      TCTTACCTGC CGTAAAATGC TCAACCCTGT GTCCTGGGCC TGGGCCTGCT GTGACTGACC  2880
      TACAGTGGAC TTTCTCTCTG GAATGGAACC TTCTTAGGCC TCCTGGTGCA ACTTAATTTT  2940
      TTTTTTTAAT GCTATCTTCA AAACGTTAGA GAAAGTTCTT CAAAAGTGCA GCCCAGAGCT  3000
      GCTGGGCCCA CTGGCCGTCC TGCATTTCTG GTTTCCAGAC CCCAATGCCT CCCATTCGGA  3060
  10  TGGATCTCTG CGTTTTTATA CTGAGTGTGC CTAGGTTGCC CCTTATTTTT TATTTTCCCT  3120
      GTTGCGTTGC TATAGATGAA GGGTGAGGAC AATCGTGTAT ATGTACTAGA ACTTTTTTAT  3180
      TAAAGAAACT TTTCCCAGAA AAAAA
```

Seq ID NO: 338 Protein sequence
Protein Accession #: NP_001784.2

```
   1          11         21         31         41         51
   |          |          |          |          |          |
  MGLPRGPLAS LLLLQVCWLQ CAASEPCRAV FREAEVTLEA GGAEQEPGQA LGKVFMGCPG   60
  QEPALFSTDN DDFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG  120
  KGPFPQRLNQ LKSNKDRDTK IFYSITGPGA DSPPEGVFAV EKETGWLLIN KPLDREEIAK  180
  YELFGHAVSE NGASVEDPMN ISIIVTDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT  240
  DEDDAIYTYN GVVAYSIHSQ EPKDPHDLMF TIHRSTGTIS VISSGLDREK VPEYTLTIQA  300
  TDMDGDGSTT TAVAVVEILD ANDNAPMFDP QKYEAHVPEN AVGHEVQRLT VTDLDAPNSP  360
  AWRATYLIMG GDDGDHFTIT THPESNQGIL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL  420
  PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGIPTGEPVC VYTAEDPDKE NQKISYRILR  480
  DPAGWLAMDP DSGQVTAVGT LDREDEQFVR NNIYEVMVLA MDNGSPPTTG TGTLLLTLID  540
  VNDREGPVPEP RQITICNQSP VRQVLNITDK DLSPHTSPFQ AQLTDDSDIY WTAEVNREGD  600
  TVVLSLKKFL KQDTYDVHLS LSDHGNKEQL TVIRATVCDC HGHVETCPGP WKGGPILPVL  660
  GAVLALLFLL LVLLLLVRKK RKIKEPLLLP EDDTRDNVFY YGEEGGGEED QDYDITQLHR  720
  GLEARPEVVL RNDVAPTIIP TPMYRPRPAN FDEIGNPIIE NLKAANTDPT APPYDTLLVF  780
  DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD
```

Seq ID NO: 339 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..672

```
   1          11         21         31         41         51
   |          |          |          |          |          |
  ATGAGGCTCC AAAGACCCCG ACAGGCCCCG GCGGGTGGGA GGCGCGCGCC CCGGGGCGGG   60
  CGGGGGCTCCC CCTACCGGCC AGACCCGGGG AGAGGCGCGC GGAGGCTGCG AAGGTTCCAG  120
  AAGGGCGGGG AGGGGGCGCC GCGCGCTGAC CCTCCCTGGG CACCGCTGGG GACGATGGCG  180
  CTGCTCGCCT TGCTGCTGGT CGTGGCCCTA CCGCGGGTGT GGACAGACGC CAACCTGACT  240
  GCGAGACAAC GAGATCCAGA GGACTCCCAG CGAACGGACG AGGGTGACAA TAGAGTGTGG  300
  TGTCATGTTT GTGAGAGAGA AAACACTTTC GAGTGCCAGA ACCCAAGGAG GTGCAAATGG  360
  ACAGAGCCAT ACTGCGTTAT AGCGGCCGTG AAAATATTTC CACGTTTTTT CATGGTTGCG  420
  AAGCAGTGCT CCGCTGGTTG TGCAGCGATG GAGAGACCCA AGCCAGAGGA GAAGCGGTTT  480
  CTCCTGGAAG AGCCCATGCC CTTCTTTTAC CTCAAGTGTT GTAAAATTCG CTACTGCAAT  540
  TTAGAGGGGC CACCTATCAA CTCATCAGTG TTCAAAGAAT ATGCTGGGAG CATGGGTGAG  600
  AGCTGTGGTG GGCTGTGGCT GGCCATCCTC CTGCTGCTGG CCTCCATTGC AGCCGGCCTC  660
  AGCCTGTCTT GA
```

Seq ID NO: 340 Protein sequence
Protein Accession #: Eos sequence

```
   1          11         21         31         41         51
   |          |          |          |          |          |
  MRLQRPRQAP AGGRRAPRGG RGSPYRPDPG RGARRLRRFQ KGGEGAPRAD PPWAPLGTMA   60
  LLALLLVVAL PRVWTDANLT ARQRDPEDSQ RTDEGDNRVW CHVCERENTF ECQNPRRCKW  120
  TEPYCVIAAV KIFPRPFMVA KQCSAGCAAM ERPKPEEKRF LLEEPMPFFY LKCCKIRYCN  180
  LEGPPINSSV FKEYAGSMGE SCGGLWLAIL LLLASIAAGL SLS
```

Seq ID NO: 341 DNA sequence
Nucleic Acid Accession #: XM_035292.2
Coding sequence: 53..1576

```
   1          11         21         31         41         51
   |          |          |          |          |          |
  GCTCGCTGGG CCGCGGCTCC CGGGTGTCCC AGGCCCGGCC GGTGCGCAGA GCATGGCGGG   60
  TGCGGGCCCG AAGCGGCGCG CGCTAGCGGC GCCGGCGGCG GAGGAGAAGG AAGAGGGCGCG  120
  GGAGAAGATG CTGGCCGCCA AGAGCGCGGA CGGCTCGGCG CCGGCAGGCG AGGGCGAGGG  180
  CGTGACCCTG CAGCGGAACA TCACGCTGCT CAACGGCGTG GCCATCATCG TGGGGACCAT  240
  TATCGGCTGG GGCATCTTCG TGACGCCCAC GGGCGTGCTG AAGGAGGCAG GCTCGCCGGG  300
  GCTGGCGCTG GTGGTGTGGG CCGCGTGCGC CGTCTTCTCC ATCGTGGGCG CGCTCTGCTA  360
  CGCGGAGCTC GGCACCACCA TCTCCAAATC GGGCGGCTAC TACGCCTACA TGCTGGAGGT  420
  CTACGGCTCG CTGCCCGGCT TCCTCAAGCT CTGGATCGAG CTGCTCATCA TCCGGACCTTC  480
  ATCGCAGTAC ATCGTGGCCC TGGTCTTCGC CACCTACCTG CTCAAGCCGC TCTTCCCCAC  540
  CTGCCCGGTG CCCGAGGAGG CAGCCAAGCT CGTGGCCTGC CTCTGCGTGC TGCTGCTCAC  600
  GGCCGTGAAC TGCTACAGCG TGAAGGCCGC CACCCGGGTC CAGGATGCCT TTGCCGCCGC  660
  CAAGCTCCTG GCCCTGGCCC TGATCATCCT GCTGGGGCTT CGTCCAGATCG GAAAGGGTGA  720
  TGTGTCCAAT CTAGATCCCA ACTTCTCATT TGAAGGCACC AAACTGGATG TGGGGAACAT  780
  TGTGCTGGCA TTATACAGCG GCCTCTTTGC CTATGGAGGA TGGAATTACT TGAATTTCGT  840
  CACAGAGGAA ATGATCAACC.CCTACAGAAA CCTGCCCCTG GCCATCATCA TCTCCCTGCC  900
  CATCGTGACG CTGGTGTACG TGCTGACCAA CCTGGCCTAC TTCACCACCC TGTCCACCGA  960
  GCAGATGCTG TCGTCCGAGG CCGTGGCCGT GGACTTCGGG AACTATCACC TGGGCGGTCAT 1020
```

```
GTCCTGGATC ATCCCCGTCT TCGTGGGCCT GTCCTGCTTC GGCTCCGTCA ATGGGTCCCT 1080
GTTCACATCC TCCAGGCTCT TCTTCGTGGG GTCCCGGGAA GGCCACCTGC CCTCCATCCT 1140
CTCCATGATC CACCCACAGC TCCTCACCCC CGTGCCGTCC CTCGTGTTCA CGTGTGTGAT 1200
GACGCTGCTC TACGCCTTCT CCAAGGACAT CTTCTCCGTC ATCAACTTCT TCAGCTTCTT 1260
CAACTGGCTC TGCGTGGCCC TGGCCATCAT CGGCATGATC TGGCTGCGCC ACAGAAAGCC 1320
TGAGCTTGAG CGGCCCATCA AGGTGAACCT GGCCCTGCCT GTGTTCTTCA TCCTGGCCTG 1380
CCTCTTCCTG ATCGCCGTCT CCTTCTGGAA GACACCCGTG GAGTGTGGCA TCGGCTTCAC 1440
CATCATCCTC AGCGGGCTGC CGGTCTACTT CTTCGGGGTG TGGTGGAAAA ACAAGCCCAA 1500
GTGGCTCCTC CAGGGCATCT TCTCCACGAC CGTCCTGTGT CAGAAGCTCA TGCAGGTGGT 1560
CCCCCAGGAG ACATAGCCAG GAGGCCGAGT GGCTGCCGGA GGAGCATGC
```

Seq ID NO: 342 Protein sequence
Protein Accession #: XP_035292.2

```
1          11         21         31         41         51
|          |          |          |          |          |
MAGAGPKRRA LAAPAABERE BAREKMLAAK SADGSAPAGE GSGVTLQRNI TLLNGVAIIV  60
GTIIGSGIFV TPTGVLKEAG SPGLALVVWA ACGVFSIVGA LCYAELGTTI SKSGGDYAYM  120
LEVYGSLPAP LKLWIELLII RPSSQYIVAL VFATYLLKPL FPTCPVPEEA AKLVACLCVL  180
LLTAVNCYSV KAATRVQDAF AAAKLLALAL IILLGFVQIG KGDVSNLDPN FSFEGTKLDV  240
GNIVLALYSG LFAYGGWNYL NFVTEEMINP YRNLPLAIII SLPIVTLVYV LTNLAYFTTL  300
STEQMLSSEA VAVDFGNYHL GVMSWIIPVF VGLSCFGSVN GSLFTSSRLF FVGSREGHLP  360
SILSMIHPQL LTPVPSLVFT CVMTLLYAFS KDIFSVINPF SFFNWLCVAL AIIGMIWLRH  420
RKPELERPIK VNLALFVPFI LACLFLIAVS FWKTPVECGI GFTIILSGLP VYFFGVWWKN  480
KPKWLLQGIF STTVLCQKLM QVVPQBT
```

Seq ID NO: 343 DNA sequence
Nucleic Acid Accession #: NM_005268.1
Coding sequence: 168..989

```
1          11         21         31         41         51
|          |          |          |          |          |
TAAAAAGCAA AAGAATTCGC GGCCGCGTCG ACACGGGCTT CCCCGAAAAC CTTCCCCGCT  60
TCTGGATATG AAATTCAAGC TGCTTGCTGA GTCCTATTGC CGGCTGCTGG GAGCCAGGAG  120
AGCCCTGAGG AGTAGTCACT CAGTAGCAGC TGACGCGTGG GTCCACCATG AACTGGAGTA  180
TCTTTGAGGG ACTCCTGAGT GGGGTCAACA AGTACTCCAC AGCCTTTGGG CGCATCTGGC  240
TGTCTCTGGT CTTCATCTTC CGCGTGCTGG TGTACCTGGT GACGGCCGAG CGTGTGTGGA  300
GTGATGACCA CAAGGACTTC GACTGCAATA CTCGCCAGCC CGGCTGCTCC AACGTCTGCT  360
TTGATGAGTT CTTCCCTGTG TCCCATGTGC GCCTCTGGGC CCTGCAGCTT ATCCTGGTGA  420
CATGCCCCTC ACTGCTCGTG GTCATGCACG TGGCCTACCG GGAGGTTCAG GAGAAGAGGC  480
ACCGAGAAGC CCATGGGGAG AACAGTGGGC GCCTCTACCT GAACCCCGGC AAGAAGCGGG  540
GTGGGCTCTG TGGACATAT GTCTGCAGCC TAGTGTTCAA GGCGAGCGTG GACATCGCCT  600
TTCTCTATGT GTTCCACTCA TTCTACCCCA AATATATCCT CCCTCCTGTG GTCAAGTGCC  660
ACGCAGATCC ATGTCCCAAT ATAGTGGACT GCTTCATCTC CAAGCCCTCA GAGAAGAACA  720
TTTTCACCCT CTTCATGGTG GCCACAGCTG CCATCTGCAT CCTGCTCAAC CTCGTGGAGC  780
TCATCTACCT GGTGAGCAAG AGATGCCACG AGTGCCTGGC AGCAAGGAAA GCTCAAGCCA  840
TGTGCACAGG TCATCACCCC CACGGTACCA CCTCTTCCTG CAAACAAGAC GACCTCCTTT  900
CGGGGTGACCT CATCTTTCTG GGCTCAGACA GTCATCCTCC TCTCTTACCA GACCGCCCCC  960
GAGACCATGT GAAGAAAACC ATCTTGTGAG GGGCTGCCTG GACTGGTCTG GCAGGTTGGG  1020
CCTGGATGGG GAGGCTCTAG CATCTCTCAT AGGTGCAACC TGAGAGTGGG GGAGCTAAGC  1080
CATGAGGTAG GGGCAGGCAA GAGAGAGGAT TCAGACGCTC TGGGAGCCAG TTCCTAGTCC  1140
TCAACTTCCAG CCACCTGCCC CAGCTCGACG GCACTGGGCC AGTTCCCCCT CTGCTCTGCA  1200
GCTCGGTTTC CTTTTCTAGA ATGGAAATAG TGAGGGGCAA TGC
```

Seq ID NO: 344 Protein sequence
Protein Accession #: NP_005259.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MNWSIFEGLL SGVNKYSTAF GRIWLSLVFI FRVLVYLVTA ERVWSDDHKD FDCNTRQPGC  60
SNVCFDEPFP VSHVRLWALQ LILVTCPSLL VVMHVAYREV QERHREAHG ENSGRLYLNP  120
GKKRGGLWWT YVCSLVFKAS VDIAPLYVFH SFYPKYILPP VVKCHADPCP NIVDCFISKP  180
SEKNIPTLFM VATAAICILL NLVELIYLVS KRCHECLAAR KAQAMCTGHH PHGTTSSCKQ  240
DDLLSGDLIF LGSDSHPPLL PDRPRDHVKK TIL
```

Seq ID NO: 345 DNA sequence
Nucleic Acid Accession #: NM_002391.1
Coding sequence: 26..457

```
1          11         21         31         41         51
|          |          |          |          |          |
CGGGCGAAGC AGCGCGGGCA GCGAGATGCA GCACCGAGGC TTCCTCCTCC TCACCCTCCT  60
CGGCCCTGCTG CGCGCTCACCT CCGCCGGTCGC CAAAAAGAAA GATAAGGTGA AGAAGGGCGG  120
CCCCGGGGAGC GAGTGCGCTG AGTGGGGCTG GGGGCCCTGC ACCCCCAGCA CGAAGGATTG  180
CGGCGTGGGT TTCCGCGAGG GCACCTGCGG GGCCCAGACC CAGCGCATCC GGTGCAGGGT  240
GCCCTGCAAC TGGAAGAAGG AGTTTGGAGC CGACTGCAAG TACAAGTTTG AGAACTGGGG  300
TGCGTGTGAT GGGGGCACAG GCACCAAAGT CCGCCAAGGC ACCCTGAAGA AGGCGCGCTA  360
CAATGCTCAG TGCCAGGAGA CCATCCGCGT CACCAAGCCC TGCACCCCCA AGACCAAAGC  420
AAAGGCCAAA GCCAAGAAAG GGAAGGGAAA GGACTAGACG CCAAGCCTGG ATGCCAAGGA  480
GCCCCTGGTG TCACATGGGG CCTGGCCACG CCCTCCCTCT CCCAGGCCCG AGATGTGACC  540
CACCAGTGCC TTCTGTCTGC TCGTTAGCTT TAATCAATCA TGCCCTGCCT TGTCCCTCTC  600
ACTCCCCAGC CCCACCCCTA AGTGCCCAAA GTGGGGAGGG ACAAGGGATT CTGGGAAGCT  660
TGAGCCTCCC CCAAAGCAAT GTGAGTCCCA GAGCCCGCTT TTGTTCTTCC CCACAATTCC  720
ATTACTAAGA AACACATCAA ATAAACTGAC TTTTTCCCCC CAATAAAAGC TCTTCTTTTT  780
```

```
TAATAT

Seq ID NO: 346 Protein sequence
Protein Accession #: NP_002382.1

1          11         21         31         41         51
|          |          |          |          |          |
MQHRGFLLLT LLALLALTSA VAKKKDKVKK GGPGSECAEW AWGPCTPSSK DCGVGFPEGT   60
CGAQTQRIRC RVPCNWKKEF GADCKYKFEN WGACDGGTGT KVRQGTLKKA RYNAQCQETI  120
RVTKPCTPKT KAKAKAKKGK GKD

Seq ID NO: 347 DNA sequence
Nucleic Acid Accession #: NM_006783.1
Coding sequence: 1..786

1          11         21         31         41         51
|          |          |          |          |          |
ATGGATTGGG GGACGCTGCA CACTTTCATC GGGGGTGTCA ACAAACACTC CACCAGCATC   60
GGGAAGGTGT GGATCACAGT CATCTTTATT TTCCGAGTCA TGATCCTAGT GGTGGCTGCC  120
CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTCGTCTGCA ACACACTGCA ACCGGGATGC  180
AAAAATGTGT GCTATGACCA CTTTTTCCCG GTGTCCCACA TCCGGCTGTG GGCCCTCCAG  240
CTGATCTTCG TCTCCACCCC AGCGCTGCTG GTGGCCATGC ATGTGGCCTA CTACAGGCAC  300
GAAACCACTC GCAAGTTCAG GCGAGGAGAG AAGAGGAATG ATTTCAAAGA CATAGAGGAC  360
ATTAAAAAGC ACAAGGTTCG GATAGAGGGG TCGCTGTGGT GGACGTACAC CAGCAGCATC  420
TTTTTCCGAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTACTTCCT TTACAATGGG   480
TACCACCTGC CCTGGGTGTT GAAATGTGGG ATTGACCCCT GCCCCAACCT TGTTGACTGC  540
TTTATTTCTA GGCCAACAGA GAAGACCGTG TTTACCATTT TTATGATTTC TGCGTCTGTG  600
ATTTGCATGC TGCTTAACGT GGCAGAGTTG TGCTACCTGC TGCTGAAAGT GTGTTTTAGG  660
AGATCAAAGA GAGCACAGAC GCAAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG  720
CAGAATGAAA TGAATGAGCT GATTTCAGAT AGTGGTCAAA ATGCAATCAC AGGTTTCCCA  780
AGCTAA

Seq ID NO: 348 Protein sequence
Protein Accession #: NP_006774.1

1          11         21         31         41         51
|          |          |          |          |          |
MDWGTLHTFI GGVNKHSTSI GKVWITVIFI FRVMILVVAA QEVWGDEQED FVCNTLQPGC   60
KNVCYDHFFP VSHIRLWALQ LIFVSTPALL VAMHVAYYRH ETTRKFRRGE KRNDFKDIED  120
IKKHKVRIEG SLWWTYTSSI FFRIIFEAAF MYVFYFLYNG YHLPWVLKCG IDPCPNLVDC  180
FISRPTEKTV FTIFMISASV ICMLLNVAEL CYLLLRVCFR RSKRAQTQKN HPNHALKESK  240
QNEMNELISD SGQNAITGFP S

Seq ID NO: 349 DNA sequence
Nucleic Acid Accession #: NM_002571.1
Coding sequence: 99..587

1          11         21         31         41         51
|          |          |          |          |          |
CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCGC AGCCATGCTG TGCCTCCTGC   60
TCACCCTGGG CGTGGCCCTG GTCTGTGGTG TCCCGGCCAT GGACATCCCC CAGACCAAGC  120
AGGACCTGGA GCTCCCAAAG TTGGCAGGGA CCTGGCACTC CATGGCCATG GCGACCAACA  180
ACATCTCCCT CATGGCGACA CTGAAGGCCC TCTGAGGGT CCACATCACC TCACTGTTGC  240
CCACCCCCGA GGACAACCTG GAGATCGTTC TGCACAGATG GGAGAACAAC AGCTGTGTTG  300
AGAAGAAGGT CCTTGGAGAG AAGACTGGGA ATCCAAAGAA GTTCAAGATC AACTATACGG  360
TGGCGAACGA GGCCACGCTG CTCGATACTG ACTACGACAA TTTCCTGTTT CTCTGCCTAC  420
AGGACACCAC CACCCCCATC CAGAGCATGA TGTGCCAGTA CCTGGCCAGA GTCCTGGTGG  480
AGGACGATGA GATCATGCAG GGATTCATCA GGGCTTTCAG GCCCCTGCCC AGGCACCTAT  540
GGTACTTGCT GGACTTGAAA CAGATGGAAG AGCCGTGCCG TTTCTAGCTC ACCTCCGCCT  600
CCAGGAAGAC CAGACTCCCA CCCTTCCACA CCTCCAGAGC AGTGGGACTT CCTCCTGCCC  660
TTTCAAAGAA TAACCACAGC TCAGAAGACG ATGACGTGGT CATCTGTGTC GCCATCCCCT  720
TCCTGCTGCA CACCTGCACC ATTGCCATGG GGAGGCTGCT CCCTGGGGGC AGAGTCTCTG  780
GCAGAGGTTA TTAATAAACC CTTGGAGCAT G

Seq ID NO: 350 Protein sequence
Protein Accession #: NP_002562.1

1          11         21         31         41         51
|          |          |          |          |          |
MDIPQTKQDL ELPKLAGTWH SMAMATNNIS LMATLKAPLR VHITSLLPTP EDNLEIVLHR   60
WENNSCVEKK VLGEKTGNPK KFKINYTVAN EATLLDTDYD NFLFLCLQDT TTPIQSMMCQ  120
YLARVLVEDD EIMQGFIRAF RPLPRHLWYL LDLKQMEEPC RF

Seq ID NO: 351 DNA sequence
Nucleic Acid Accession #: NM_006500.1
Coding sequence: 27..1967

1          11         21         31         41         51
|          |          |          |          |          |
ACTTGCGTCT CGCCCTCCGG CCAAGCATGG GGCTTCCCAG GCTGGTCTGC GCCTTCTTGC   60
TCGCCGCCTG CTGCTGCTGT CCTCGCGTCG CGGGTGTGCC CGGAGAGGCT GAGCAGCCTG  120
CGCCTGAGCT GGTGGAGGTG GAAGTGGGCA GCACAGCCCT TCTGAAGTGC GGCCTCTCCC  180
AGTCCCAAGG CAACCTCAGC CATGTCGACT GGTTTTCTGT CCACAAGGAG AAGCGGACGC  240
TCATCTTCCG TGTGCGCCAG GGCCAGGGCC AGAGCGAACC TGGGGAGTAC GAGCAGCGGC  300
```

```
TCAGCCTCCA GGACAGAGGG GCTACTCTGG CCCTGACTCA AGTCACCCCC CAAGACGAGC    360
GCATCTTCTT GTGCCAGGGC AAGCGCCCTC GGTCCCAGGA GTACCGCATC CAGCTCCGCG    420
TCTACAAAGC TCCGGAGGAG CCAAACATCC AGGTCAACCC CCTGGGCATC CCTGTGAACA    480
GTAAGGAGCC TGAGGAGGTC GCTACCTGTG TAGGGAGGA CGGGTACCCC ATTCCTCAAG    540
TCATCTGGTA CAAGAATGGC CGGCCTCTGA AGGAGGAGAA GAACCGGGTC CACATTCAGT    600
CGTCCCAGAC TGTGGAGTCG AGTGGTTTGT ACACCTTGCA GAGTATTCTG AAGGCACAGC    660
TGGTTAAAGA AGACAAAGAT GCCCAGTTTT ACTGTGAGCT CAACTACCGG CTGCCCAGTG    720
GGAACCACAT GAAGGAGTCC AGGGAAGTCA CCGTCCCTGT TTTCTACCCG ACAGAAAAAG    780
TGTGGCTGGA AGTGGAGCCC GTGGGAATGC TGAAGGAAGG GGACCGCGTG GAAATCAGGT    840
GTTTGGCTGA TGGCAACCCT CCACCACACT TCAGCATCAG CAAGCAGAAC CCCAGCACCA    900
GGGAGGCAGA GGAAGAGACA ACCAACGACA ACGGGGTCCT GGTGCTGGAG CCTGCCCGGA    960
AGGAACACAG TGGGCGCTAT GAATGTCAGG CCTGGAACTT GGACACCATG ATATCGCTGC   1020
TGAGTGAACC ACAGGAACTA CTGGTGAACT ATGTGTCTGA CGTCCGAGTG AGTCCCGCAG   1080
CCCCTGAGAG ACAGGAAGGC AGCAGCCTCA CCCTGACCTG TGAGGCAGAG AGTAGCCAGG   1140
ACCTCGAGTT CCAGTGGCTG AGAGAAGAGA CAGACCAGGT GCTGGAAAGG GGGCCTGTGC   1200
TTCAGTTGCA TGACCTGAAA CGGGAGGCAG GAGGCGGCTA TCGCTGCGTG GCGTCTGTGC   1260
CCAGCATACC CGGCCTGAAC CGCACACAGC TGGTCAAGCT GGCCATTTTT GGCCCCCCTT   1320
GGATGGCATT CAAGGAGAGG AAGGTGTGGG TGAAAGAGAA TATGGTGTTG AATCTGTCTT   1380
GTGAAGCGTC AGGGCACCCC CGGCCCACCA TCTCCTGGAA CGTCAACGGC ACGGCAAGTG   1440
AACAAGACCA AGATCCACAG CGAGTCCTGA GCACCCTGAA TGTCCTCGTG ACCCCGGAGC   1500
TGTTGGAGAC AGGTGTTGAA TGCACGGCCT CCAACGACCT GGGCAAAAAC ACCAGCATCC   1560
TCTTCCTGGA GCTGGTCAAT TTAACCACCC TCACACCAGA CTCCAACACA ACCACTGGCC   1620
TCAGCACTTC CACTGCCAGT CCTCATACCA GAGCCAACAG CACCTCCACA GAGAGAAAGC   1680
TGCCGGAGCC GGAGAGCCGG GGCGTGGTCA TCGTGGCTGT GATTGTGTGC ATCCTGGTCC   1740
TGGCGGTGCT GGGCGCTGTC CTCTATTTCC TCTATAAGAA GGGCAAGCTG CCGTGCAGGC   1800
GCTCAGGGAA GCAGGAGATC ACGCTGCCCC GTCTCGTAA GACCGAACTT GTAGTTGAAG   1860
TTAAGTCAGA TAAGCTCCCA GAAGAGATGG GCCTCCTGCA GGGTAGCAGC GGTGACAAGA   1920
GGGCTCCGGG AGACCAGGGA GAGAAATACA TCGATCTGAG GCATTAGCCC CGAATCACTT   1980
CAGCTCCCTT CCCTGCCTGG ACCATTCCCA GCTCCCTGCT CACTCTTCTC TCAGCCAAAG   2040
CCTCCAAAGG GACTAGAGAG AAGCCTCCTG CTCCCCTCAC CTGCACACCC CCTTTCAGAG   2100
GGCCACTGGG TTAGGACCTG AGGACCTCAC TTGGCCCTGC AAGCCGCTTT TCAGGGACCA   2160
GTCCACCACC ATCTCCTCCA CGTTGAGTGA AGCTCATCCC AAGCAAGGAG CCCCAGTCTC   2220
CCGAGCGGGT AGGAGAGTTT CTTGCAGAAC GTGTTTTTTC TTTACACACA TTATGGCTGT   2280
AAATACCTGG CTCCTGCCAG CAGCTGAGCT GGGTAGCCTC TCTGAGCTGG TTTCCTGCCC   2340
CAAAGGCTGG CTTCCACCAT CCAGGTGCAC CACTGAAGTG AGTACACACC GGAGCCAGGC   2400
GCCTGCTCAT GTTGAAGTGC GCTGTTCACA CCCGCTCCGG AGAGCACCCC AGCGGCATCC   2460
AGAAGCAGCT GCAGTGTTGC TGCCACCACC CTCCTGCTCG CCTCTTCAAA GTCTCCTGTG   2520
ACATTTTTTC TTTGGTCAGA AGCCAGGAAC TGGTGTCATT CCTTAAAAGA TACGTGCCGG   2580
GGCCAGGTGT GGTGGCTCAC GCCTGTAATC CCAGCACTTT GGGAGGCCGA GGCGGGCGGA   2640
TCACAAAGTC AGGACGAGAC CATCCTGGCT AACACGGTGA AACCCTGTCT CTACTAAAAA   2700
TACAAAAAAA AATTAGCTAG GCGTAGTGGT GGGCACCTAT AGTCCCAGCT ACTCGGAAGG   2760
CTGAAGCAGG AGAATGGTAT GAATCCAGGA GGTGGAGCTT GCAGTGAGCC GAGACCGTGC   2820
CACTGCACTC CAGCCTGGGC AACACAGCGA GACTCCGTCT CGAGGAAAAA AAAAGAAAAG   2880
ACGCGTACCT GCGGTGAGGA AGCTGGGCGC TGTTTTCGAG TTCAGGTGAA TTAGCCTCAA   2940
TCCCCGTGTT CACTTGCTCC CATAGCCCTC TTGATGGATC ACGTAAAACT GAAAGGCAGC   3000
GGGGAGCACA CAAAGATGAG GTCTACACTG TCCTTCATTA GGATTAAAGC TATGGTTATA   3060
TTAGCACCAA ACTTCTACAA ACCAAGCTCA GGGCCCCAAC CCTAGAAGGG CCCAAATGAG   3120
AGAATGGTAC TTAGGGATGG AAAACGGGGC CTGGCTAGAG CTTCGGGTGT GTGTGTCTGT   3180
CTGTGTGTAT GCATACATAT GTGTGTATAT ATGGTTTTGT CAGGTGTGTA AATTTGCAAA   3240
TTGTTTCCTT TATATATGTA TGTATATATA TATATGAAAA TATATATATA TATGAAAAAT   3300
AAAGCTTAAT TGTCCCAGAA AATCATACAT TGCTTTTTTA TTCTACATGG GTACCACAGG   3360
AACCTGGGGG CCTGTGAAAC TACAACCAAA AGGCACACAA AACGTTTCC AGTTGGCAGG   3420
AGAGATCAGG GGTTACCTCT GCTTCTGAGC AAATGGCTCA AGCTCTACCA GAGCAGACAG   3480
CTACCCTACT TTTCAGCAGC AAAACGTCCC GTATGACGCA GCACGAAGGG CCTGGCAGGC   3540
TGTTAGCAGG AGCTATGTCC CTTCCTATCG TTTCCGTCCA CTT
```

Seq ID NO: 352 Protein sequence
Protein Accession #: NP_006491.1

```
1         11        21        31        41        51
|         |         |         |         |         |
GLPRLVCAFL LAACCCCPRV AGVPGEABQP APELVEVEVG STALLKIGLS QSQGNLSHVD    60
WFSVEKEKRT LIFRVRQGQG QSEPGEYEQR LSLQDRGATL ALTQVTPQDE RIFLCQGKRP   120
RSQEYRIQLR VYKAPEEPNI QVNPLGIPVN SKEPEEVATC VGRNGYPIPQ VIWYKNGRPL   180
KEEKNRVHIQ SSQTVESSGL YTLQSILEAQ LVKEDKDAQF YCKINYRLPS GNHMKESREV   240
TVPVFYPTEK VWLEVEPVGM LKEGDRVEIR CLADGNPPPH FSISKQNPST REAEEETTND   300
NGVLVLEPAR KEHSGRYECQ AWNLDTMISL LSEPQELLVN YVSDVRVSPA APERQEGSSL   360
TLTCEAESSQ DLEFQWLREE TDQVLERGPV LQLHDLKREA GGYRCVASV PSIPGLNRTQ   420
LVKLAIFGPP WMAFKERKVW VKENMVLNLS CEASGHPRPT ISWNVNGTAS EQDQDPQRVL   480
STLNVLVTPE LLETGVECTA SNDLGKNTSI LFLELVNLTT LTPDSNTTTG LSTSTASPHT   540
RANSTSTERK LPEPESRGVV IVAVIVCILV LAVLGAVLYF LYKKGKLPCR RSGKQEITLP   600
PSRKTELVVE VKSDKLPEEM GLLQGSSGDK RAPGDQGEKY IDLRH
```

Seq ID NO: 353 DNA sequence
Nucleic Acid Accession #: NM_003183.3
Coding sequence: 165..2639

```
1         11        21        31        41        51
|         |         |         |         |         |
TCGAGCCTGG CGGTAGAATC TTCCCAGTAG GCGGCGCGGG AGGGAAAAGA GGATTGAGGG    60
GCTAGGCCGG GCGGATCCCG TCCTCCCCCG ATGTGAGCAG TTTTCCGAAA CCCCGTCAGG   120
CGAAGGCTGC CCAGAGAGGT GGAGTCGGTA GCGGGGCCGG GAACATGAGG CAGTCTCTCC   180
TATTCCTGAC CAGCGTGGTT CCTTTCGTGC TGGCGCCGCG ACCTCCGGAT GACCCGGGCT   240
TCGGCCCCCA CCAGAGACTC GAGAAGCTTG ATTCTTTGCT CTCAGACTAC GATATTCTCT   300
```

```
CTTTATCTAA  TATCCAGCAG  CATTCGGTAA  GAAAAAGAGA  TCTACAGACT  TCAACACATG   360
TAGAAACACT  ACTAACTTTT  TCAGCTTTGA  AAAGGCATTT  TAAATTATAC  CTGACATCAA   420
GTACTGAACG  TTTTTCACAA  AATTTCAAGG  TCGTGGTGGT  GGATGGTAAA  AACGAAAGCG   480
AGTACACTGC  AAAATGGCAG  GACTTCTTCA  CTGGACACGT  GGTTGGTGAG  CCTGACTCTA   540
GGGTTCTAGC  CCACATAAGA  GATGATGATG  TTATAATCAG  AATCAACACA  GATGGGGCCG   600
AATATAACAT  AGAGCCACTT  TGGAGATTTG  TTAATGATAC  CAAAGACAAA  AGAATGTTAG   660
TTTATAAATC  TGAAGATATC  AAGAATGTTT  CACGTTTGCA  GTCTCCAAAA  GTGTGTGGTT   720
ATTTAAAAGT  GGATAATGAA  GAGTTGCTCC  CAAAAGGGTT  AGTAGACAGA  GAACCACCTG   780
AAGAGCTTGT  TCATCGAGTG  AAAAGAAGAG  CTGACCCAGA  TCCCATGAAG  AACACGTGTA   840
AATTATTGGT  GGTAGCAGAT  CATCGCTTCT  ACAGATACAT  GGGCAGAGGG  GAAGAGAGTA   900
CAACTACAAA  TTACTTAATA  GAGCTAATTG  ACAGAGTTGA  TGACATCTAT  CGGAACACTT   960
CATGGGATAA  TGCAGGTTTT  AAAGGCTATG  GAATACAGAT  AGAGCAGATT  CGCATTCTCA  1020
AGTCTCCACA  AGAGGTAAAA  CCTGGTGAAA  AGCACTACAA  CATGGCAAAA  AGTTACCCAA  1080
ATGAAGAAAA  GGATGCTTGG  GATGTGAAGA  TGTTGCTAGA  GCAATTTAGC  TTTGATATAG  1140
CTGAGGAAGC  ATCTAAAGTT  TGCTTGGCAC  ACCTTTTCAC  ATACCAAGAT  TTTGATATGG  1200
GAACTCTTGG  ATTAGCTTAT  GTTGGCTCTC  CCAGAGCAAA  CAGCCATGGA  GGTGTTTGTC  1260
CAAAGGCTTA  TTATAGCCCA  GTTGGGAAGA  AAAATATCTA  TTTGAATAGT  GGTTTGACGA  1320
GCACAAAGAT  TTATGGTAAA  ACCATCCTTA  CAAAGGAAGC  TGACCTGGTT  ACAACTCATG  1380
AATTGGGACA  TAATTTTGGA  GCAGAACATG  ATCCGGATGG  TCTAGCAGAA  TGTGCCCCGA  1440
ATGAGGACCA  GGGAGGGAAA  TATGTCATGT  ATCCCATAGC  TGTGAGTGGC  GATCACGAGA  1500
ACAATAAGAT  GTTTTCAAAC  TGCAGTAAAC  AATCAATCTA  TAAGACCATT  GAAAGTAAGG  1560
CCCAGGAGTG  TTTTCAAGAA  CGCAGCAATA  AAGTTTGTGG  GAACTCGAGG  GTGGATGAAG  1620
GAGAAGAGTG  TGATCCTGGC  ATCATGTATC  TGAACAACGA  CACCTGCTGC  AACAGCGACT  1680
GCACGTTGAA  GGAAGGTGTC  CAGTGCAGTG  ACAGGAACAG  TCCTTGCTGT  AAAAACTGTC  1740
AGTTTGAGAC  TGCCCAGAAG  AAGTGCCAGG  AGGCGATTAA  TGCTACTTGC  AAAGGCGTGT  1800
CCTACTGCAC  AGGTAATAGC  AGTGAGTGCC  CGCCTCCAGG  AAATGCTGAA  AATGACACTG  1860
TTTGCTTGGA  TCTTGGCAAG  TGTAAGGATG  GGAAATGCAT  CCCTTTCTGC  GAGAGGGAAC  1920
AGCAGCTGGA  GTCCTGTGCA  TGTAATGAAA  CTGACAACTC  CTGCAAGGTG  TGCTGCAGGG  1980
ACCTTTCTGG  CCGCTGTGTG  CCCTATGTCG  ATGCTGAACA  AAAGAACTTA  TTTTTGAGGA  2040
AAGGAAAGCC  CTGTACAGTA  GGATTTTGTG  ACATGAATGG  CAAATGTGAG  AAACGAGTAC  2100
AGGATGTAAT  TGAACGATTT  TGGGATTTCA  TTGACCAGCT  GAGCATCAAT  ACTTTTGGAA  2160
AGTTTTTAGC  AGACAACATC  GTTGGGTCTG  TCCTGGTTTT  CTCCTTGATA  TTTTGGATTC  2220
CTTTCAGCAT  TCTTGTCCAT  TGTGTGGATA  AGAAATTGGA  TAAACAGTAT  GAATCTCTGT  2280
CTCTGTTTCA  CCCCAGTAAC  GTCGAAATGC  TGAGCAGCAT  GGATTCTGCA  TCGGTTCGCA  2340
TTATCAAACC  CTTTCCTGCG  CCCCAGACTC  CAGGCCGCCT  GCAGCCTGCC  CCTGTGATCC  2400
CTTCGGCGCC  AGCAGCTCCA  AAACTGGACC  ACCAGAGAAT  GGACACCATC  CAGGAGACC   2460
CCAGCACAGA  CTCCCATATG  GACGAGGATG  GGTTTGAGAA  GGACCCCTTC  CCAAATAGCA  2520
GCACAGCTGC  CAAGTCATTT  GAGGATCTCA  CGGACCATCC  GGTCGCCAGA  AGTGAAAAGG  2580
CTGCCTCCTT  TAAAACTGCAG  CGTCAGAATC  GTGTTAACAG  CAAAGAAACA  GAGTGCTAAT  2640
TTAGTTCTCA  GCTCTTCTGA  CTTAAGTGTG  CAAAATATTT  TTATAGATTT  GACCTACAAA  2700
TCAATCACAG  CTTGTATTTT  GTGAAGACTG  GGAAGTGACT  TAGCAGATGC  TGGTCATGTG  2760
TTTGAACTTC  CTGCAGGTAA  ACAGTTCTTG  TGTGGTTTGG  CCCTTCTCCT  TTTGAAAAGG  2820
TAAGGTGAAA  GTGAATCTAC  TTATTTTGAG  GCTTTCAGGT  TTTAGTTTTT  AAAATATCTT  2880
TTGACCTGTG  GTGCAAAAGC  AGAAAATACA  GCTGGATTGG  GTTATGAATA  TTTACGTTTT  2940
TGTAAATTAA  TCTTTTATAT  TGATAACAGC  ACTGACTAGG  GAAATGATCA  GTTTTTTTTT  3000
ATACACTGTA  ATGAACGCT   GAATATGAAG  CATTTGGCAT  TTATTTGTGA  GAAAAGTGGA  3060
ATAGTTTTTT  TTTTTTTTTT  TTTTTTTTGC  CTTCAACTAA  AAACAAAGGA  GATAAATTTA  3120
GTATACATTG  TATCTAAATT  GTGGGTCTAT  TTCTAGTTAT  TACCCAGAGT  TTTTATGTAG  3180
CAGGGAAATA  ATATATCTAA  ATTTAGAAAT  CATTTGGGTT  AATATGGCTC  TTCATAATTC  3240
TAAGACTAAT  GCTCAGAACC  TAACCACTAC  CTTACAGTGA  GGGCTATACA  TGGTAGCCAG  3300
TTGAATTTAT  GGAATCTACC  AACTGTTTAG  GGCCCTGATT  TGCTGGGCAG  TTTTTCTGTA  3360
TTTTATAAGT  ATCTTCATGT  ATCCCTGTTA  CTGATAGGGA  TACATGTCTT  AGAAAATTCA  3420
CTATTGGCTG  GGAGTGGTGG  CTCATGCCTG  TAATCCCAGC  ACTTGGAGAG  GCTGAGGTTG  3480
CGCCACTACA  CTCCAGCCTG  GGTGACAGAG  TGAGATCTGC  CTC
```

Seq ID NO: 354 Protein sequence
Protein Accession #: NP_003174.2

```
        1         11        21        31        41        51
        |         |         |         |         |         |
    MRQSLLFLTS  VVPFVLAPRP  PDDPGFGPHQ  RLEKLDSLLS  DYDILSLSNI  QQHSVRKRDL    60
    QTSTHVETLL  TFSALKRHFK  LYLTSSTERF  SQNFKVVVVD  GKNRSEYTAK  WQDFFTGHVV   120
    GEPDSRVLAH  IRDDDVIIRI  NTDGAEYNIE  PLWRFVNDTK  DKRMLVYKSE  DIKNVSRLQS   180
    PKVCGYLKVD  NEELLPKGLV  DREPPEELVH  RVKRRADPDP  MKNTCKLLVV  ADHRFYRMG    240
    RGEESTTNY   LIELIDRVDD  IYRNTSWDNA  GFKGYGIQIE  QIRILKBPQE  VKPGEKHYNM   300
    AKSYPNEEKD  AWDVKMLLEQ  FSFDIAEEAS  KVCLAHLFTY  QDFDMGTLGL  AYVGSPRANS   360
    HGGVCPKAYY  SPVGKKNIYL  NSGLTSTKNY  GKTILTKEAD  LVTTHELGHN  FGAEHDPDGL   420
    AECAPNEDQG  GKYVMYPIAV  SGDHENNKMF  SNCSKQSIYK  TIESKAQECF  QERSNKVCGN   480
    SEVDEGEECD  PGIMYLNNDT  CCNSDCTLKE  GVQCSDRNSP  CCKNCQFETA  QKKCQEAINA   540
    TCKGVSYCTG  NSSECPPPGN  AENDTVCLDL  GKCKDGKCIP  FCEREQQLES  CACNETDNSC   600
    KVCCRDLSGR  CVPYVDAEQK  NLPLRKGKPC  TVGFCDMNGK  CEKRVQDVIE  RFWDFIDQLS   660
    INTPGKFLAD  NIVGSVLVFS  LIFWIPFSIL  VHCVDKKLDK  QYESLSLFHP  SNVEMLSSMD   720
    SASVRIIKPF  PAPQTPGRLQ  PAPVIPSAPA  APKLDHQRMD  TIQEDPSTDS  HMDEDGFEKD   780
    PFPNSSTAAK  SFEDLTDHPV  ARSEKAASFK  LQRQNRVNSK  ETEC
```

Seq ID NO: 355 DNA sequence
Nucleic Acid Accession #: NM_021832.1
Coding sequence: 164..2248

```
        1         11        21        31        41        51
        |         |         |         |         |         |
    TCGAGCCTGG  CGGTAGAATC  TTCCCAGTAG  GCGGCGCGGG  AGGAAAAGAG  GATTGAGGGG    60
    CTAGCCGGG   CGGATCCCGT  CCTCCCCCGA  TGTGAGCAGT  TTTCCGAAAC  CCCGTCAGGC   120
    GAAGGCTGCC  CAGAGAGGTG  GAGTCGGTAG  CGGGGCCGGG  AACATGAGGC  AGTCTCTCCT   180
```

```
      ATTCCTGACC AGCGTGGTTC CTTTCGTGCT GGCGCCGCGA CCTCCGGATG ACCCGGGCTT    240
      CGGCCCCCAC CAGAGACTCG AGAAGCTTGA TTCTTTGCTC TCAGACTACG ATATTCTCTC      300
      TTTATCTAAT ATCCAGCAGC ATTCGGTAAG AAAAAGAGAT CTACAGACTT CAACACATGT      360
 5    AGAAACACTA CTAACTTTTT CAGCTTTGAA AAGGCATTTT AAATTATACC TGACATCAAG       420
      TACTGAACGT TTTTCACAAA ATTTCAAGGT CGTGGTGGTG GATGGTAAAA ACGAAAGCGA      480
      GTACACTGTA AAATGGCAGG ACTTCTTCAC TGGACACGTG GTTGGTGAGC CTGACTCTAG      540
      GGTTCTAGCC CACATAAGAG ATGATGATGT TATAATCAGA ATCAACACAG ATGGGGCCGA      600
      ATATAACATA GAGCCCACTT GGAGATTTGT TAATGATACC AAAGACAAAA GAATGTTAGT      660
      TTATAAATCT GAAGATATCA AGAATGTTTC ACGTTTGCAG TCTCCAAAAG TGTGTGGTTA      720
 10   TTTAAAAGTG GATAATGAAG AGTTGCTCCC AAAAGGGTTA GTAGACAGAG AACCACCTGA      780
      AGAGCTTGTT CATCGAGTGA AAAGAAGAGC TGACCCAGAT CCCATGAAGA ACACGTGTAA       840
      ATTATTGGTG GTAGCAGATC ATCGCTTCTA CAGATACATG GGCAGAGGGG AAGAGAGTAC      900
      AACTACAAAT TACTTAATAG AGCTAATTGA CAGAGTTGAT GACATCTATC GGAACACTTC      960
      ATGGGATAAT GCAGGTTTTA AAGGCTATGG AATACAGATA GAGCAGATTC GCATTCTCAA     1020
 15   GTCTCCACAA GAGGGTAAAA CCTGGTGAAA GCACTACAAC ATGGCAAAAA GTTACCCAAA     1080
      TGAAGAAAAG GATGCTTGGG ATGTGAAGAT GTTGCTAGAG CAATTTAGCT TTGATATAGC     1140
      TGAGGAAGCA TCTAAAGTTT GCTTGGCACA CCTTTTCACA TACCAAGATT TTGATATGGG     1200
      AACTCTTGGA TTAGCTTATG TTGGCTCTCC CAGAGCAAAC AGCCATGGAG GTGTTTGTCC     1260
      AAAGGCTTAT TATAGCCCAG TTGGGAAGAA AAATATCTAT TTGAATAGTG GTTTGACGAG     1320
 20   CACAAAGAAT TATGGTAAAA CCATCCTTAC AAAGGAAGCT GACCTGGTTA CAACTCATGA     1380
      ATTGGGACAT AATTTTGGAG CAGAACATGA TCCGGATGGT CTAGCAGAAT GTGCCCCGAA     1440
      TGAGGACCAG GGAGGGAAAT ATGTCATGTA TCCCATAGCT GTGAGTGGCG ATCACGAGAA     1500
      CAATAAGATG TTTTCAAACT GCAGTAAACA ATCAATCTAT AAGACCATTG AAAGTAAGGC     1560
      CCAGGAGTGT TTTCAAGAAC GCAGCAATAA AGTTTGTGGG AACTCGAGGG TGGATGAAGG     1620
 25   AGAAGAGTGT GATCCTGGCA TCATGTATCT GAACAACGAC ACCTGCTGCA ACAGCGACTG     1680
      CACGTTGAAG GAAGGTGTCC AGTGCAGTGA CAGGAACAGT CCTTGCTGTA AAAACTGTCA     1740
      GTTTGAGACT GCCCAGAAGA AGTGCCAGGA GGCGATTAAT GCTACTTGCA AAGGCGTGTC     1800
      CTACTGCACA GGTAATAGCA GTGAGTGCCC GCCTCCAGGA AATGCTGAAG ATGACACTGT     1860
      TTGCTTGGAT CTTGGCAAGT GTAAGGATGG GAAATGCATC CCTTTCTGCG AGAGGGAACA     1920
 30   GCAGCTGGAG TCCTGTGCAT GTAATGAAAC TGACAACTCC TGCAAGGTGT GCTGCAGGGA     1980
      AGGAAAGCCC TGTACAGTAG GATTTTGTGA CATGAATGGC AAATGTGAGA AACGAGTACA     2040
      GGATGTAATT GAACGATTTT GGGGATTTCAT TGACCAGCTG AGCATCAATA CTTTTTGGAAA    2100 (?)
```

Seq ID NO: 356 Protein sequence
Protein Accession #: NP_068604.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MRQSLLFLTS VVPFVLAPRF PDDPGFGPHQ RLEKLDSLLS DYDILSLSNI QQHSVRKRDL   60
QTSTHVETLL TFSALKRHFK LYLTSSTERF SQNFKVVVVD GKNESEYTVK WQDFFTGHVV  120
GEPDSRVLAH IRDDDVIIRI NTDGAEYNIE PLWRFVNDTK DKRMLVYKSE DIKNVSRLQS  180
PKVCGYLKVD NEELLPKGLV DREPPEELVE RVKRRADPDP MKNTCKLLVV ADHRFYRYMG  240
RGEESTTTNY LIELIDRVDD IYRNTSWDNA GFKGYGIQIE QIRILKSPQE VKPGEKHYNM  300
AKSYPNEEKD AWDVKMLLEQ FSFDIAEEAS KVCLAHLFTY QDFDMGTLGL AYVGSPRANS  360
HGGVCPKAYY SPVGKKNIYL NSGLTSTKNY GKTILTKEAD LVTTHELGHN FGAEHDPDGL  420
AKCAPNEDQG GKYVMYPIAV SGDHENNKMF SNCSKQSIYK TIESKAQECF QERSNKVCGN  480
SRVDEGEECD PGIMYLNNDT CCNSDCTLKE GVQCSDRNSP CCKNCQFETA QKKCQEAINA  540
TCKGVSYCTG NSSECPPPGN AEDDTVCLDL GKCKDGKCIP FCEREQQLES CACNETDNSC  600
KVCCRDLSGR CVPYVDAEQK NLFLRKGKPC TVGFCDMNGK CEKRVQDVIE RFWDFIDQLS  660
INTFGKFLAD NIVGSVLVFS LIFWIPFSIL VHCV
```

Seq ID NO: 357 DNA sequence
Nucleic Acid Accession #: NM_004994.1
Coding sequence: 20..2143

```
1          11         21         31         41         51
|          |          |          |          |          |
AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCCTG GTCCTGGTGC TCCTGGTGCT   60
GGGCTGCTGC TTTGCTGCCC CCAGACAGCG CCAGTCCACC CTTGTGCTCT TCCCTGGAGA  120
CCTGAGAACC AATCTCACCG ACAGGCAGCT GGCAGAGGAA TACCTGTACC GCTATGGTTA  180
CACTCGGGTG GCAGAGATGC GTGGAGAGTC GAAATCTCTG GGGCCTGCGC TGCTGCTTCT  240
```

1089

```
      CCAGAAGCAA CTGTCCCTGC CCGAGACCGG TGAGCTGGAT AGCGCCACGC TGAAGGCCAT   300
      GCGAACCCCA CGGTGCGGGG TCCCAGACCT GGGCAGATTC CAAACCTTTG AGGGCGACCT   360
      CAAGTGGCAC CACCACAACA TCACCTATTG GATCCAAAAC TACTCGGAAG ACTTGCCGCG   420
 5    GGCGGTGAATT GACGACGCCT TTGCCCGCGG CTTCGCACTG TGGACGCGCT TGACGCCGCT   480
      CACCTTCACT CGCGTGTACA GCCGGGACGC AGACATCGTC ATCCAGTTTG GTGTCGCGGA   540
      GCACGGAGAC GGGTATCCCT TCGACGGGAA GGACGGGCTC CTGGCACACG CCTTTCCTCC   600
      TGGCCCGGGC ATTCAGGGAG ACGCCCATTT CGACGATGAC GAGTTGTGGT CCCTGGGCAA   660
      GGGCGTCGTG GTTCCAACTG GGTTTGGAAA CGCAGATGGC GCGGCCTGCC ACTTCCCCTT   720
 10   CATCTTCGAG GGCCGCTCCT ACTCTGCCTG CACCACCGAC GGTCGCTCCG ACGGCTTGCC   780
      CTGGTGCAGT ACCACGGCCA ACTACGACAC CGACGACCGG TTTGGCTTCT GCCCCAGCGA   840
      GAGACTCTAC ACCCGGGACG GCAATGCTGA TGGGAAACCC TGCCAGTTTC CATTCATCTT   900
      CCAAGGCCAA TCCTACTCCG CCTGCACCAC GGACGGTCGC TCCGACGGCT ACCGCTGGTG   960
      CGCCACCACC GCCAACTACG ACCGGGACAA GCTCTTCGGC TTCTGCCCGA CCCGAGCTGA  1020
 15   CTCGACGGTG ATGGGGGGCA ACTCGGCGGG GGAGCTGTGC GTCTTCCCCT TCACTTTCCT  1080
      GGGTAAGGAG TACTCGACCT GTACCAGCGA GGGCCGCGGA GATGGGCGCC TCTGGTGCGC  1140
      TACCACCTCG AACTTTGACA GCGACAAGAA GTGGGGCTTC TGCCCCGGACC AAGGATACAG  1200
      TTTGTTCCTC GTGGCGGCGC ATGAGTTCGG CCACGCGCTG GCCTTAGATC ATTCCTCAGT  1260
      GCCGGAGGCG CTCATGTACC CTATGTACCG CTTCACTGAG GGGCCCCCCT TGCATAAGGA  1320
 20   CGACGTGAAT GGCATCCGGC ACCTCTATGG TCCTCGCCCT GAACCTGAGC CACGGCCTCC  1380
      AACCACCACC ACACCGGCAGC CCACGGCTCC CCGGACGGTC TGCCCCACCG GACCCCCCAC  1440
      TGTCCACCCC TCAGAGCGCC CCACAGCTGG CCCCACAGGT CCCCCCTCAG CTGGCCCCAC  1500
      AGGTCCCCCC ACTGCTGGCC CTTCTACGGC CACTACTGTG CCTTTGAGTC CGGTGGACGA  1560
      TGCCTGCAAC GTGAACATCT TCGACGCCAT CGCGGAGATC GGGAACCAGC TGTATTTGTT  1620
      CAAGGATGGG AAGTACTGGC GATTCTCTGA GGGCAGGGGG AGCCGGCCGC AGGGCCCCTT  1680
 25   CCTTATCGCC GACAAGTGGC CCGCGCTGCC CCGCAAGCTG GACTCCGTCT TTGAGGAGCC  1740
      GCTCTCCAAG AAGCTTTTCT TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGCGTC  1800
      GGTGCTGGGC CCGAGGCGTC TGGACAAGCT GGGCCTGGGA GCCGACGTGG CCCAGGTGAC  1860
      CGGGGCCCTC CGGAGTGGCA GGGGGAAGAT GCTGCTGTTC AGCGGGCGGC GCCTCTGGAG  1920
      GTTCGACGTG AAGGCGCAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGGATGTT  1980
 30   CCCCGGGGTG CCTTTGGACA CGCACGACGT CTTCCAGTAC CGAGAGAAAG CCTATTTCTG  2040
      CCAGGACCGC TTCTACTGGC GCGTGAGTTC CCGGAGTGAG TTGAACCAGG TGGACCAAGT  2100
      GGGCTACGTG ACCTATGACA TCCTGCAGTG CCCTGAGGAC TAGGGCTCCC GTCCTGCTTT  2160
      GCAGTGCCAT GTAAATCCCC ACTGGGACCA ACCCTGGGGA AGGACCCCAGT TTGCCGGATA  2220
      CAAACTGGTA TTCTGTTCTG GAGGAAGGGG AGGAGTGGAG GTGGGCTGGG CCCTCTCTTC  2280
 35   TCACCTTTGT TTTTTGTTGG AGTGTTTCTA ATAAACTTGG ATTCTCTAAC CTTT
```

Seq ID NO: 358 <u>Protein sequence</u>
Protein Accession #: NP_004985.1

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       MSLWQPLVLV LLVLGCCPAA PRQRQSTLVL FPGDLRTNLT DRQLAEEYLY RYGYTRVAEM    60
       RGESKSLGPA LLLLQKQLSL PETGELDSAT LKAMRTPRCG VPDLGRFQTF EGDLKWHHHN   120
       ITYWIQNYSE DLPRAVIDDA FARAFALWSA VTPLTFTRVY SRDADIVIQF GVAEHGDGYP   180
       FDGKDGLLAH AFPPGPGIQG DAHFDDDELW SLGKGVVVPT RFGNADGAAC HFPFIFEGRS   240
       YSACTTDGRS DGLPWCSTTA NYDTDDRFGF CPSERLYTRD GNADGKPCQF PFIFQGQSYS   300
       ACTTDGRSDG YRWCATTANY DRDKLFGFCP TRADSTVMGG NSAGELCVFP FTFLGKEYST   360
       CTSEGRGDGR LWCATTSNFD SDKKWGFCPD QGYSLFLVAA HEFGHALGLD HSSVPEALMY   420
       PMYRFTEGPP LHKDDVNGIR HLYGPRPEPE PRPPTTTTPQ PTAPPTVCPT GPPTVHPSER   480
       PTAGPTGPPS AGPTGPPTAG PSTATTVPLS PVDDACNVNI FDAIAEIGNQ LYLFKDGKYW   540
       RFSEGRGSRP QGPFLIADKW PALPRKLDSV FEEPLSKKLF FFSGRQVWVY TGASVLGPRR   600
       LDKLGLGADV AQVTGALRSG RGKMLLFSGR RLWRFDVKAQ MVDPRSASEV DRMFPGVPLD   660
       THDVFQYREK AYFCQDRFYW RVSSRSELNQ VDQVGYVTYD ILQCPED
```

Seq ID NO: 359 <u>DNA sequence</u>
Nucleic Acid Accession #: NM_000213.1
Coding sequence: 127..5385

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       CGCCCGCBCG CTGCAGCCCC ATCTCCTAGC GGCAGCCCAG GCGCGGAGG AGCGAGTCCG    60
       CCCCGAGGTA GGTCCAGGAC GGGCGCACAG CAGCAGCCGA GGCTGGCCGG GAGAGGGAGG   120
       AAGAGGATGG CAGGGCCACG CCCCAGCCCA TGGGCCAGGC TGCTCCTGGC AGCCTTGATC   180
       AGCGGTCAGCC TCTCTGGGAC CTTGGCAAAC CGCTGCAAGA AGGCCCCAGT GAAGAGCTGC   240
       ACGGAGTGTG TCCGTGTGGA TAAGGACTGC GCCTACTGCA CAGACGAGAT GTTCAGGGAC   300
       CGGCGCTGCA ACACCCAGGC GGAGCTGCTG GCCGCGGGCT GCCAGCGGGA GAGCATCGTG   360
       GTCATGGAGA GCAGCTTCCA AATCACAGAG GAGACCCAGA TTGACACCAC CCTGCGCGAC   420
       AGCCAGATGT CCCCCCAAGG CCTGCGGGTC CGTCTGCGGC CCGGTGAGGA GCGGCATTTT   480
       GAGCTGGAGG TGTTTGAGCC ACTGGAGAGC CCCGTGGACC TGTACATCCT CATGGACTTC   540
       TCCAACTCCA TGTCCGATGA TCTGGACAAC CTCAAGAAGA TGGGGCAGAA CCTGGCTCGG   600
       GTCCTGAGCG AGCTCACCAG CGACTACACT ATTGGATTTG GCAAGTTTGT GGACAAAGTC   660
       AGCGTCCCGC AGACGGACAT GAGGCCTGAG AAGCTGAAGG AGCCCTGGCC CAACAGTGAC   720
       CCCCCCTTCT CCTTCAAGAA CGTCATCAGC CTGACAGAAG ATGTGGATGA GTTCCGGAAT   780
       AAACTGCAGG GAGAGCGGAT CTCAGGCAAC CTGGATGCTC CTGAGGGCGG CTTCGATGCC   840
       ATCCTGCAGA CAGCTGTGTG CACGGAGGAC ATTGGCTGGC GCCGGACAG CACCCACCTG   900
       CTGGTCTTCT CCACCGAGTC AGCCTTCCAC TATGAGGCTG ATGGCGCCAA CGTGCTGGCT   960
       GGCATCATCA GCCGCAACGA TGAACGGTGC CACCTGGACA CCACGGGCAC CTACACCCAG  1020
       TACAGGACAC AGGACTACCC GTCGGTGCCC ACCCTGGTGC GCCTGCTCGC CAAGCACAAC  1080
       ATCATCCCCA TCTTTGCTGT CACCAACTAC TCCTATAGCT ACTACGAGAA GCTTCACACC  1140
       TATTTCCCTG TCTCCTCACT GGGGGTGCTG CAGGAGGACT CGTCCAACAT CGTGGAGCTG  1200
       CTGGAGGAGG CCTTCAATCG GATCCGCTCC AACCTGGACA TCCGGGCCCT AGACAGCCCC  1260
       CGAGGCCTTC GGACAGAGGT CACCTCCAAG ATGTTCCAGA AGACCAGGAC TGGGTCCTTT  1320
       CACATCCGGC GGGGGGAAGT GGGTATATAC CAGGTGCAGC TGCGGGCCCT TGAGCACGTG  1380
       GATGGGACGC ACGTGTGCCA GCTGCCGGAG GACCAGAAGG GCAACATCCA TCTGAAACCT  1440
```

```
     TCCTTCTCCG ACGGCCTCAA GATGGACGCG GGCATCATCT GTGATGTGTG CACCTGCGAG  1500
  5  CTGCAAAAAG AGGTGCGGTC AGCTCGCTGC AGCTTCAACG GAGACTTCGT GTGCGGACAG  1560
     TGTGTGTGCA GCGAGGGCTG GAGTGGCCAG ACCTGCAACT GCTCCACCGG CTCTCTGAGT  1620
     GACATTCAGC CCTGCCTGCG GGAGGGCGAG GACAAGCCGT GCTCCGGCCG TGGGGAGTGC  1680
     CAGTGCGGGC ACTGTGTGTG CTACGGCGAA GGCCGCTACG AGGGTCAGTT CTGCGAGTAT  1740
     GACAACTTCC AGTGTCCCCG CACTTCCGGG TTCCTCTGCA ATGACCGAGG ACGCTGCTCC  1800
     ATGGGCCAGT GTGTGTGTGA GCCTGGTTGG ACAGGCCCAA GCTGTGACTG TCCCCTCAGC  1860
     AATGCCACCT GCATCGACAG CAATGGGGCG ATCTGTAATG GACGTGGCCA CTGTGAGTGT  1920
 10  GGCCGCTGCC ACTGCCACCA GCAGTGCCTC TACACGGACA CCATCTGCGA GATCAACTAC  1980
     TCGGCGATCC ACCCGGGCCT CTGCGAGGAC CTACGCTCCT GCGTGCAGTG CCAGGCGTGG  2040
     GGCACCGGCG AGAAGAAGGG GCGCACGTGT GAGGAATGCA ACTTCAAGGT CAAGATGGTG  2100
     GACGAGCTTA AGAGAGCCGA GGAGGTGGTG GTGCGCTGCT CCTTCCGGGA CGAGGATGAC  2160
 15  GACTGCACCT ACAGCTACAC CATGGAAGGT GACGGCGCCC CTGGGCCCAA CAGCACTGTC  2220
     CTGGTGCACA AGAAGAAGGA CTGCCCCTCCG GGCTCCTTCT GGTGGCTCAT CCCCCTGCTC  2280
     CTCCTCCTCC TGCCGCTCCT GGCCCTGCTA CTGCTGCTAT GCTGAAGTA CTGTGCCTGC  2340
     TGCAAGGCCT GCCTGGCACT TCTCCCGTGC TGCAACCGAG GTCACATGGT GGGCTTTAAG  2400
     GAAGACCACT ACATGCTGCG GGAGAACCTG ATGGCCTCTG ACCACTTGGA CACGCCCATG  2460
 20  CTGCGCAGCG GGAACCTCAA GGGGCCGTGAC GTGGTCCGCT GGAAGGTCAC CAACAACATG  2520
     CAGCGGCCTG GCTTTGCCAC TCATGCCGCC AGCATCAACC CCACAGAGCT GGTGCCCTAC  2580
     GGGCTGTCCT TGCGCCTGGC CCGCCTTTGC ACCGAGAACC TGCTGAAGCC TGACACTCGG  2640
     GAGTGCGCCC AGCTGCGCCA GGAGGTGGAG GAGAACCTGA ACGAGGTCTA CAGGCAGATC  2700
 25  TCCGGTGTAC ACAAGCTCCA GCAGACCAAG TTCCGGCAGC AGCCCAATGC CGGGAAAAAG  2760
     CAAGACCACA CCATTGTGGA CACAGTGCTG ATGGCGCCCC GCTCGGCCAA GCCGGCCCTG  2820
     CTGAAGCTTA CAGAGAAGCA GGTGGAACAG AGGGCCTTCC ACGACCTCAA GGTGGCCCCC  2880
     GGCTACTACA CCCTCACTGC AGACCAGGAC GCCCGGGGCA TGGTGGAGTT CCAGGAGGGC  2940
     GTGGGAGCTGG TGGACGTACG GGTGCCCCTC TTTATCCGGC CTGAGGATGA CGACGAGAAG  3000
 30  CAGCTGCTGG TGGAGGCCAT CGACGTGCCC GCAGGCACTG CCACCCTCGG CCGCCGCCTG  3060
     GTAAACATCA CCATCATCAA GGAGCAAGCC AGAGACGTGG TGTCCTTTGA GCAGCCTGAG  3120
     TTCTCGGTCA GCCGCGGGGA CCAGGTGGCC CGCATCCCTG TCATCCGGCG TGTCCTGGAC  3180
     GGGCGGGAAGT CCCAGGTCTC CTACCGCACA CAGGATGGCG CCGCGCAGGG CAACCCGGAC  3240
 35  TACATCCCCG TGGAGGGTGA GCTGCTGTTC CAGCCTGGGG AGGCCTGGAA AGAGCTGCAG  3300
     GTGAAGCTCC TGGAGCTGCA AGAAGTTGAC TCCCTCCTGC GGGGCCGGCCA GGTCCGCCGT  3360
     TTCCACAGTCC AGCTCAGCAA CCCTAAGTTT GGGGCCCAACC TGGGCCAGCC CCACTCCACC  3420
     ACCATCATCA TCAGGGACCC AGATGAACTG GACCGGGAGC TCACGAGTCA GATGTTGTCA  3480
     TCACAGCCAC CCCCTCACGG CGACCTGGGC GCCCCGCAGA ACCCCAATGC TAAGGCCGCT  3540
 40  GGGTCCAGGA AGATCCATTT CAACTGGCTG CCCCCTTCTG GCAAGCCAAT GGGGTACAGG  3600
     GTAAAGTACT GGATTCAGGG TGACTCCGAA TCCGAAGCCC ACCTGCTCGA CAGCAAGGTG  3660
     CCCCTCAGTGG AGCTCACCAA CCTGTACCCG TATTGCGACT ATGAGATGAA GGTGTGCGCC  3720
     TACGGGGCTC AGGGCGAGGG ACCCTACAGC TCCCTGGTGT CCTGCCGCAC CCACCAGGAA  3780
     GTGCCCAGCG AGGCAGGGCG TCTGGCCTTC AATGTCGTCT CCTCCACGGT GACCCAGCTG  3840
     AGCTGGGCTG AGCCGGCTGA GACCAACGGT GAGATCACAC CCTACGAGGT CTGCTATGGC  3900
     CTGGTCAACG ATGACAACCG ACCTATTGGG CCCATGAAGA AAGTGCTGGT TGACAACCCT  3960
     AAGAACCGGA TGCTGCTTAT TGAGAACCTT CGGGAGTCCC AGCCCTACCG CTACACGGTG  4020
 45  AAGGCGCGCA ACGGGGCCGG CTGGGGGGCCT GAGCGGGAGG CCATCATCAA CCTGGCCACC  4080
     CAGCCCAAGA GGCCCATGTC CATCCCCATC ATCCCTGACA TCCCTATCGT GGACGCCCAG  4140
     AGCGGGGAGG ACTACGACAG CTTCCTTATG TACAGCGATG ACGTTCTACG CTCTCCATCG  4200
     GGCAGCCAGA GGCCCAGCGT CTCCGATGAC ACTGAGCACC TGGTGAATGG CCGGATGGAC  4260
     TTTGCCTTCC CGGGCAGCAC CAACTCCCTG CACAGGATGA CCACGACCAG TGCTGCTGCC  4320
 50  TATGGCCACC ACCTGAGCCC ACACGTGCCC CACCGCGTGG TAAGCACATC CTCCACCCTC  4380
     ACACGGGACT ACAACTCACT GACCCGCTCA GAACACTCAC ACTCGACCAC ACTGCCGAGG  4440
     GACTACTCCA CCCTCACCTC CGTCTCCTCC CACGACTCTC GCCTGACTGC TGGTGTGCCC  4500
     GACACGCCCA CCCGCCTGGT GTTCTCTGCC CTGGGGCCCA CATCTCTCAG AGTGAGCTGG  4560
     CAGGAGCCGC GGTGCGAGCG GCCGCTGCAG GGCTACAGTG TGGAGTACCA GCTGCTGAAC  4620
 55  GGCGGTGAGC TGCATCGGCT CAACATCCCC AACCCTGCCC AGACCTGGT GGTGGTGGAA  4680
     GACCTCCTGC CCAACCACTC CTACGTGTTC CGCGTGCGGG CCCAGAGCCA GGAAGGCTGG  4740
     GGCCGAGAGC GTGAGGGTGT CATCACCATT GAATCCCAGG TGCACCCGCA GAGCCCACTG  4800
     TGTCCCCTGC CAGGCTCCGC CTTCACTTTG AGCACTCCCA GTGCCCCAGG CCCGCTGGTG  4860
     TTCACTGCCC TGAGCCCAGA CTCGCTGCAG CTGAGCTGGG AGCGGCCACG GAGGCCCAAT  4920
     GGGGATATCG TCGGCTACCT GGTGACCTGT GAGATGGCCAC AAGGAGGAGG GCCAGCCACC  4980
 60  GCATTCGGGG TGGATGGAGA CAGCCCCGAG AGCCGGCTGA CCGTGCCGGG CCTCAGCGAG  5040
     AACGTGCCCT ACAAGTTCAA GGTGCAGGCC AGGACCACTG AGGGCTTCGG GCCAGAGCGC  5100
     GAGGGCATCA TCACCATAGA GTCCCAGGAT GGAGGACCCT TCCCGCAGCT GGGGCAGCCGT  5160
     GCCGGGCTCT TCCAGCACCC GCTGCAAAGC GAGTACAGCA GCATCACCAC CACCCACACC  5220
 65  AGCGCCACCG AGCCCTTCCT AGTGGATGGG CCGACCCTGG GGCCCAGCA CCTGGAGGCA  5280
     GGCGGCTCCC TCACCCGGCA TGTGACCCAG GAGTTTGTGA GCCGGACACT GACCACCAGC  5340
     GGAACCCTTA GCACCCACAT GGACCAACAG TTCTTCCAAA CTTGACCGCA CCCTGCCCCA  5400
     CCCCCGCCAT GTCCCACTAG GCGTCCTCCC GACTCCTCTC CCGGAGCCTC CTCAGCTACT  5460
     CCATCCTTGC ACCCCTGGGG CCCCAGCCCA CCGGCATGCA GCTAGGTGTC  5520
 70  TCCTGGGAGG CATGAAGGGG GCAAGGTCCG TCCTCTGTGG GCCCAAACCT ATTTGTAACC  5580
     AAAGAGCTGG GAGCAGCACA AGGACCCAGC CTTTGTTCTG CACTTAATAA ATGGTTTTGC  5640
     TACTG
```

Seq ID NO: 360 **Protein sequence**
Protein Accession #: NP_000204.1

```
     1          11         21         31         41         51
     |          |          |          |          |          |

     MAGPRPSPWA RLLLAALISV SLSGTLANRC KKAPVKSCTE CVRVDKDCAY CTDEMFRDRR   60
     CNTQAELLAA GCQRESIVVM ESSFQITEET QIDTTLRRSQ MSPQGLRVRL RPGEERHPEL  120
     EVFEPLESFV DLYILMDFSN SMSDDLDNLK KMGQNLARVL SQLTSDYTIG FGKFVDKVSV  180
     PQTDMRPEKL KEPWPNSDPP FSFKNVISLT EDVDEFRNKL QGERISGNLD APEGGFDAIL  240
     QTAVCTRDIG WRPDSTHLLV FSTESAFHYE ADGANVLAGI MSRNDERCHL DTTGTYTQYR  300
     TQDYPSVPTL VRLLAKHNII PIFAVTNYSY SYYEKLHTYF PVSSLGVLQE DSSNIVELLE  360
     EAFNRIRSNL DIRALDSPRG LRTEVTSKMF QKTRTGSFHI RRGEVGIYQV QLRALEHVDG  420
```

1091

```
       THVCQLPEDQ KGNIHLKPSF SDGLKMDAGI ICDVCTCELQ KEVRSARCSF NGDFVCGQCV    480
       CSEGWSGQTC NCSTGSLSDI QPCLREGEDK PCSGRGECQC GHCVCYGEGR YEGQFCEYDN    540
       FQCPRTSGFL CNDRGRCSMG QCVCEPGWTG PSCDCPLSNA TCIDSNGGIC NGRGHCECGR    600
       CHCHQQSLYT DTICEINYSA IHPGLCEDLR SCVQCQAWGT GEKKGRTCEE CNFKVKMVDE    660
 5     LKRAEEVVVR CSFRDEDDDC TYSYTMEGDG APGPNSTVLV HKKKDCPPGS FWWLIPLLLL    720
       LLPLLALLLL LCWKYCACCK ACLALLPCCN RGHMVGPKED HYMLRENLMA SDHLDTPMLR    780
       SGNLKGRDVV RNKVTNNMQR PGPATHAASI NPTELVPYGL SLRLARLCTE NLLKPDTREC    840
       AQLRQEVEEN LNEVYRQISG VHKLQQTKFR QQPNAGKKQD HTIVDTVLMA PRSAKPALLK    900
       LTEKQVEQRA PHDLKVAPGY YILTADQDAR GMVSFQEGVE LVDVRVPLFI RPEDDDEKQL    960
10     LVEAIDVPAG TATLGRRLVN ITIIKEQARD VVSFEQPEFS VSRGDQVARI PVIRRVLDGG   1020
       KSQVSYRTQD GTAQGNRDYI PVEGELLFOP GEAWKELQVK LLELQEVDSL LRGRQVRRFH   1080
       VQLSNPKFGA HLGQPHSTTI IIRDPDELDR SFTSQMLSSQ PPPHGDLGAP QNPMAKAAGS   1140
       RKIHPNWLPP SGKPMGYRVK YWIQGDSESE AHLLDSKVPS VELTNLYPYC DYEMKVCAYG   1200
       AQGSGPYSSL VSCRTHQEVP SEPGRLAFNV VSSTVTQLSW AEPAETNGEI TAYEVCYGLV   1260
15     NDDNRPIGPM KKVLVDNPKN RMLLIENLRE SQPYRYTVKA RNGAGWGPER EAIIHLATQP   1320
       KRPMSIPIIP DIPIVDAQSG EDYDSPLMYS DDVLRSPSGS QRPSVSDDTE HLVNGRMDFA   1380
       PPGSTNSLHR MTTTSAAAYG THLSPHVPHR VLSTSSTLTR DYNSLTRSEH SHSTTLPRDY   1440
       STLTSVSSHD SRLTAGVPDT PTRLVFSALG PTSLRVSWQE PRCERPLQGY SVEYQLLNGG   1500
       ELHRLNIPNP AQTSVVVEDL LPNHSYVFRV RAQSQEGWGR EREGVITIES QVHPQSPLCP   1560
20     LPGSAFTLST PSAPGPLVFT ALSPDSLQLS WERPRRPNGD IVGYLVTCEM AQGGGPATAF   1620
       RVDGDSPESR LTVPGLSENV PYKFKVQART TEGFGPEREG IITIESQDGG PFPQLGSRAG   1680
       LFQHPLQSEY SSITTHTSA TEPPLVDGPT LGAQHLEAGG SLTRHVTQEF VSRTLTTSGT   1740
       LSTHMDQQFF QT
```

Seq ID NO: 361 DNA sequence
Nucleic Acid Accession #: NM_013332.1
Coding sequence: 1..63

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       GCACGAGGGC GCTTTTGTCT CCGGTGAGTT TTGTGGCGGG AAGCTTCTGC GCTGGTGCTT     60
       AGTAACCGAC TTTCCTCCGG ACTCCTGCAC GACCTGCTCC TACAGCCGGC GATCCACTCC    120
       CGGCTGTTCC CCCGGAGGGT CCAGAGGGCC TTCAGAAGGA GAAGGCAGCT CTGTTTCTCT    180
       GCAGAGGAGT AGGGTCCTTT CAGCCATGAA GCATGTGTTG AACCTCTACC TGTTAGGTGT    240
       GGTACTGACC CTACTCTCCA TCTTCGTTAG AGTGATGGAG TCCCTAGAAG GCTTACTAGA    300
       GAGCCCATCG CCTGGGACCT CCTGGACCAC CAGAAGGCAG CTAGCCAACA CAGAGCCCAC    360
       CAAGGGCCTT CCAGACCATC CATCCAGAAG CATGTGATAA GACCCTCCTTC CATACTGGCC    420
       ATATTTTGGA ACACTGACCT AGACATGTCC AGATGGGAGT CCCATTCCTA GCAGACAAGC    480
       TGAGCACCGT TGTAACCAGA GAACTATTAC TAGGCCTTGA AGAACCTGTC TAACTGGATA    540
       CTCATTGCCT GGGCAAGGCC TGTTTAGGCC GGTTGCCGGTG GCTCATGCCT GTAATCCTAG    600
       CACTTTGGGA GGCTGAGGTG GGTGGATCAC CTGAGGTCAG GAGTTCGAGA CCAGCCTCGC    660
       CAACATGGCG AAACCCCATC TCTACTAAAA ATACAAAAGT TAGCTGGGTG TGGTGGCAGA    720
       GGCCTGTAAT CCCAGTTCCT TGGGAGGCTG AGGCGGGAGA ATTGCTTGAA CCCGGGGACG    780
       GAGGTTGCAG TGAACCGAGA TCGCACTGCT GTACCCAGCC TGGGCCACAG TGCAAGACTC    840
       CATCTCAAAA AAAAAAAGAA AAGAAAAAGC CTGTTTAATG CACAGGTGTG AGTGGATTGC    900
       TTATGGCTAT GAGATAGGTT GATCTCGCCC TTACCCCGGG GTCTGGTGTA TGCTGTGCTT    960
       TCCTCAGCAG TATGGCTCTG ACATCTCTTA GATGTCCCAA CTTCAGCTGT TGGGAGATGG   1020
       TGATATTTTC AACCCTACTT CCTAAACATC TGTCTGGGGT TCCTTTAGTC TTGAATGTCT   1080
       TATGCTCAAT TATTTGGTGT TGAGCCTCTC TTCCACAAGA GTTTGGGTAG   1140
       CAGTTGAAGA GGTTGTGTGG GTGGGCTGTT GGGAGTGAGG ATGGAGTGTT CAGTGCCCAT   1200
       TTCTCATTTT ACATTTTAAA GTCGTTCCTC CAACATAGTG TGTATTGGTC TGAAGGGGGT   1260
       GGTGGGATGC CAAAGCCTGC TCAAGTTATG GACATTGTGG CCACCATGTG GCTTAAATGA   1320
       TTTTTTCTAA CTAATAAAGT GGAATATATA TTTCAAAAAA AAAAAAAAAA AA
```

Seq ID NO: 362 Protein sequence
Protein Accession #: NP_037464.1

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       MKHVLNLYLL GVVLTLLSIF VRVMESLEGL LESPSPGTSW TTRSQLANTE PTKGLPDHPS     60
       RSM
```

Seq ID NO: 363 DNA sequence
Nucleic Acid Accession #: NM_023915.1
Coding sequence: 250..1326

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       GGCACGAGGG TTTCGTTTTC ATGCTTTACC AGAAAATCCA CTTCCCTGCC GACCTTAGTT     60
       TCAAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG    120
       GTGAATGGAC AGCCAGCCAC CACATGAAA GAAATCAATAAC CAGGAATAAC CTATGCTGAA    180
       CCCACGCCTC AATCGTCCCC AAGTGTTTCC TGACACGCAT CTTTGCTTAC AGTGCATCAC    240
       AACTGAAGAA TGGGGTTCAA CTTGACGCTT GCAAAATTAC CAAATAACGA GCTGCACGGC    300
       CAAGAGCATC ACAATTCAGG CAACAGGAGC GACGGGCCAG GAAAGAACAC CACCCTTCAC    360
       AATGAATTTG ACACAATTGT CTTGCCGGTG CTTTATCTCA TTATATTTGT GGCAAGCATC    420
       TTGCTGAATG GTTTAGCAGT GTGGATCTTC TTCCACATTA GGAATAAAAC CAGCTTCATA    480
       TTCTATCTCA AAAACATAGT GGTTGCAGAC CTCATAATGA CGCTGACATT TCCATTTCGA    540
       ATAGTCCATG ATGCAGGATT TGGACCTTGG TACTTCAAGT TTATTCTCTG CAGATACACT    600
       TCAGTTTTGT TTTATGCAAA CATGTATACT TCCATCGTGT TCCTTGGGCT GATAAGCATT    660
       GATCGCTATC TGAAGGTGGT CAAGCCATTT GGGGACTGTC GGATGTACAG CATAACCTTC    720
       ACGAAGGTTT TATCTGTTTG TGTTTGGGTG ATCATGGCTG TTTTGTCTTT GCCAAACATC    780
       ATCCTGACAA ATGGTCAGCC AACAGAGGAC AATATCCATG ACTGCTCAAA ACTTAAAAGT    840
       CCTTTGGGGG TCAAATGGCA TACGGCAGTC ACCTATGTGA ACAGCTGCTT GTTTGTGGCC    900
       GTGCTGGTGA TTCTGATCGG ATGTTACATA GCCTATCCA GGTACATCCA CAAATCCAGC    960
```

```
AGGCAATTCA TAAGTCAGTC AAGCCGAAAG CGAAAACATA ACCAGAGCAT CAGGGTTGTT  1020
GTGGCTGTGT TTTTTACCTG CTTTCTACCA TATCACTTGT GCAGAATTCC TTTTACTTTT  1080
AGTCACTTAG ACAGGCTTTT AGATGAATCT GCACAAAAAA TCCTATATTA CTGCAAAGAA  1140
ATTACACTTT TCTTGTCTGC GTGTAATGTT TGCCTGGATC CAATAATTTA CTTTTTCATG  1200
TGTAGGTCAT TTTCAAGAAG GCTGTTCAAA AAATCAAATA TCAGAACCAG GAGTGAAAGC  1260
ATCAGATCAC TGCAAAGTGT GAGAAGATCG GAAGTTCGCA TATATTATGA TTACACTGAT  1320
GTGTAGGCCT TTTATTGTTT GTTGGAATCG ATATGTACAA AGTGTAAATA AATGTTTCTT  1380
TTCATTATCC TTAAAAAAAA AA
```

Seq ID NO: 364 Protein sequence
Protein Accession #: NP_076404

```
1          11         21         31         41         51
|          |          |          |          |          |
MGFNLTLAKL PNNELHGQES RNSGNRSDGP GKNTTLHNEF DTIVLPVLYL IIFVASILLN   60
GLAVWIFFHI RNKTSFIFYL KNIVVKPFGDS TLTFPFRIVH DAGFGPWYFK FILCRYTSVL  120
FYANMYTSIV FLGLISIDRY LKVVKPFGDS RMYSITFTKV LSVCVWVIMA VLSLPNIILT  180
NGQPTEDNIH DCSKLKSPLG VKWHTAVTYV NSCLFVAVLV ILIGCYIAIS RYIHKSSRQF  240
ISQSSRKRKH NQSIRVVVAV FFTCFLPYHL CRIPFTFSHL DRILLDESAQK ILYYCKEITL  300
PLSACNVCLD PIIYFFMCRS PSRRLFKKSN IRTRSESIRS LQSVRRSEVR IYYDYTDV
```

Seq ID NO: 365 DNA sequence
Nucleic Acid Accession #: NM_005365.1
Coding sequence: 1..948

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGTCTCTCG AGCAGAGGAG TCCGCACTGC AAGCCTGATG AAGACCTTGA AGCCCAAGGA   60
GAGGACTTGG GCCTGATGGG TGCACAGGAA CCCACAGGCG AGGAGGAGGA GACTACCTCG  120
TCCTCTGACA GCAAGGAGGA GGAGGTGTCT GCTGCTGGGT CATCAAGTCC TCCCCAGAGT  180
CCTCAGGGAG GCGCTTCCTC CTCCATTTCC GTCTACTACA CTTTATGGAG CCAATTCGAT  240
GAGGGCTCCA GCAGTCAAGA AGAGGAAGAG CCAAGCTCCT CGGTCGACCC AGCTCAGCTG  300
GAGTTCATGT TCCAAGAAGC ACTGAAATTG AAGGTGGCTG AGTTGGTTCA TTTCCTGCTC  360
CACAAATATC GAGTCAAGGA GCCGGTCACA AAGGCAGAAA TGCTGGAGAG CGTCATCAAA  420
AATTACAAGC GCTACTTTCC TGTGATCTTC GGCAAAGCCT CCGAGTTCAT GCAGGTGATC  480
TTTGGCACTG ATGTGAAGGA GGTGGACCCC GCCGGCCACT CCTACATCCT TGTCACTGCT  540
CTTGGCTCT CGTGCGATAG CATGCTGGGT GATGGTCATA GCATGCCCAA GGCCGCCCTC  600
CTGATCATTG TCCTGGGTGT GATCCTAACC AAAGACAACT GCGCCCCTGA AGAGGTTATC  660
TGGGAAGCGT TGAGTGTGAT GGGGGTGTAT GTTGGGAAGG AGCACATGTT CTACGGGGAG  720
CCCAGGAAGC TGCTCACCCA AGATTGGGTG CAGGAAAACT ACCTGGAGTA CCGGCAGGTG  780
CCCGGCAGTG ATCCTGCGCA CTACGAGTTC CTGTGGGGTT CCAAGGCCCA CGCTGAAACC  840
AGCTATGAGA AGGTCATAAA TTATTTGGTC ATGCTCAATG CAAGAGAGCC CATCTGCTAC  900
CCATCCCTTT ATGAAGAGGT TTTGGGAGAG GAGCAAGAGG GAGTCTGA
```

Seq ID NO: 366 Protein sequence
Protein Accession #: NP_005356.1

```
1          11         21         31         41         51
|          |          |          |          |          |
MSLEQRSPHC KPDEDLEAQG EDLGLMGAQE PTGEEEETTS SSDSKEEEVS AAGSSSPPQS   60
PQGGASSSIS VYYTLWSQFD EGSSSQEEEE PSSSVDPAQL EFMFQEALKL KVAELVHFLL  120
HKYRVKEPVT KAEMLESVIK NYKRYFPVIF GKASEFMQVI FGTDVKEVDP AGHSYILVTA  180
LGLSCDSMLG DGHSMPKAAL LIIVLGVILT KDNCAPEEVI WEALSVMGVY VGKEHMFYGE  240
PRKLLTQDWV QENYLEYRQV PGSDPAHYEF LWGSKAHAET SYRKVINYLV MLNAREPICY  300
PSLYEEVLGE EQEGV
```

Seq ID NO: 367 DNA sequence
Nucleic Acid Accession #: NM_014400
Coding sequence: 86..1126

```
1          11         21         31         41         51
|          |          |          |          |          |
GGTTACTCAT CCTGGGCTCA GGTAAGAGGG CCCGAGCTCG GAGGCGGCAC ACCCAGGGGG   60
GACGGCAAGG GAGCAGGACG GAGCCATGGA CCCCGCCAAG AAAGCAGGTG CCCAGGCCAT  120
GATCTGGACT GCAGGCTGGC TGCTGCTGCT GCTGCTTCGC GGAGGAGCGC AGGCCCTGGA  180
GTGCTACAGC TGCGTGCAGA AAGCAGATGA CGGATGCTCC CCGAACAAGA TGAAGACAGT  240
GAAGTGCGCG CCGGGCGTGG ACGTCTGCAC CGAGGCGTGC ATGGCGGTGG AGACCATCCA  300
CGGACAATTC TCGCTGGCAG TGCCGGGTTG CGGTTCGGGA CTCCCCGGCA AGAATGACCG  360
CGGCCTGGAT CTTCACGGGC TTCTGGCGTT CATCCAGCTG CAGCAATGCG CTCAGGATCG  420
CTGCAACGCC AAGCTCAACC TCACCTCGCG GGCGCTCGAC CCGGCAGGTA ATGAGAGTGC  480
ATACCCGCCC AACGGCGTGG AGTGCTACAG CTGTGTGGGC CTGAGCCGGG AGGCGTGCCA  540
GGGTACAATCG CCGCCGGTCG TGAGCTGCTA CAACGCCAGC GATCATGTCT ACAAGGGCTG  600
CTTCGACGGC AACGTCACCT TGACGGCAGC TAATGTGACT GTGTCCTTGC CTGTCCGGGG  660
CTGTGTCCAG GATGAATTCT GCACTCGGGA TGGAGTAACA GGCCCAGGGT TCACGGCTCAG  720
TGGCTCCTGT TGCCAGGGGT CCCGCTGTAA CTCTGACCTC CGCAACAAGA CCTACTTCTC  780
CCCTCGAATC CCACCCCTTG TCCGGCTGCC CCCTCCAGAG CCCACGACTG TGGCCTCAAC  840
CACATCTGTC ACCACTTCTA CCTCGGCCCC AGTGAGACCC ACATCCACCA CCAAACCCAT  900
GCCAGCGCCA ACCAGTCAGA CTCCGAGACA GGGAGTAGAA CACGAGGCCT CCCGGGATGA  960
GGAGCCCAGG TTGACTGGAG GCGCCGCTGG CCACCAGGAC CGCAGCAATT CAGGGCGAGTA  1020
TCCTGCAAAA GGGGGGCCCC AGCAGCCCCA TAATAAAGGC TGTGTGGGCTC CCACAGCTGA  1080
ATTGGCAGCC CTTCTGTTGG CCGTGGCTGC TGGTGTCCTA CTGTGAGCTT CTCCACCTGG  1140
AAATTTCCCT CTCACCTACT TCTCTGGCCC TGGGTACCCC TCTTCTCATC ACTTCCTGTT  1200
CCCACCCACTG GACTGGGCTG GCCCAGCCCC TGTTTTTCCA ACATTCCCCA GTATCCCCAG  1260
CTTCTGCTGC GCTGGTTTGC GGCTTGGGA AATAAAAATAC CGTTGTATAT ATTCTGGCAG  1320
```

```
GGGTGTTCTA GCTTTTTGAG GACAGCTCCT GTATCCTTCT CATCCTTGTC TCTCCGCTTG  1380
TCCTCTTGTG ATGTTAGGAC AGAGTGAGAG AAGTCAGCTG TCACGGGGAA GGTGAGAGAG  1440
AGGATGCTAA GCTTCCTACT CACTTTCTCC TAGCCAGCCT GGACTTTGGA GCCGTGGGGTG  1500
GGTGGGACAA TGGCTCCCCA CTCTAAGCACA TGCCTCCCCT ACTCCCCGCA TCTTTGGGGA  1560
ATCGGTTCCC CATATGTCTT CCTTACTAGA CTGTGAGCTC CTCGAGGGCA GGGACCGTGC  1620
CTTATGTCTG TGTGTGATCA GTTTCTGGCA CATAAATGCC TCAATAAAGA TTTAATTACT  1680
TTGTATAGTG AAAAAAAA
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MDPARKAGAQ AMIWTAGWLL LLLLRGGAQA LECYSCVQKA DDGCSPNKMK TVKCAPGVDV   60
CTEAVGAVET IHGQFSLAVX GCGSGLPGKN DRGLDLHGLL AFIQLQQCAQ DRCNAKLNLT  120
SRALDPAGNE SAYPPNGVEC YSCVGLSREA CQGTSPPVVS CYNASDHVYK GCFDGNVTLT  180
AANVTVSLPV RGCVQDEFCT RDGVTGPGFT LSGSCCQGSR CNSDLRNKTY FSFRIPPLVR  240
LPPPEPTTVA STTSVTTSTS APVRPTSTTK PMPAPTSQTP RQGVEHEASR DEEPRLTGGA  300
AGHQDRSNSG QYPAKGGPQQ PHNKGCVAPT AGLAALLLAV AAGVLL
```

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCCGGTGC AGCTGACGAC AGCCCTGCGT GTGGTGGGCA CCAGCCTGTT TGCCCTGGCA   60
GTGCTGGGTG GCATCCTGGC AGCCTATGTG ACGGGCTACC AGTTCATCCA CACGGAAAAG  120
CACTACCTGT CCTTCGGCCT GTACGGCGCC ATCCTGGGCC TGCACCTGCT CATTCAGAGC  180
CTTTTTGCCT TCCTGGAGCA CCGGCGCATG CGACGTGCCG GCCAGGCCCT GAAGCTGCCC  240
TCCCCGCGGC GGGGGCTCGG GGGCACTGTGC ATTGCCGCAT ACCAGGAGGA CCCTGACTAC  300
TTGCGCAAGT GCCTGCGCTC GGCCCAGCGC ATCTCCTTCC CTGACCTCAA GGTGGTCATG  360
GTGGTGGATG GCAACCGCCA GGAGGACGCC TACATGCTGG ACATCTTCCA CGAGGTGCTG  420
GGCGGCACCG AGCAGGCCGG CTTCTTTGTG TGGCGCAGCA ACTTCCATGA GGCAGGCGAG  480
GGTGAGACGG AGGCCAGCCT GCAGGAGGGC ATGGACCGTG TGCGGGATGT GGTGCGGGCC  540
AGCACCTTCT CGTGCATCAT GCAGAAGTGG GGAGGCAAGC GCGAGGTCAT GTACACGGCC  600
TTCAAGGCCC TCGGCGATTC GGTGGACTAC ATCCAGGTGT GCGACTCTGA CACTGTGCTG  660
GATCCAGCCT GCACCATCGA GATGCTTCGA GTCCTGGAGG AGGATCCCCA AGTAGGGGGA  720
GTCGGGGGAG ATGTCCAGAT CCTCAACAAG TACGGACTCAT GGATTTCCTT CCTGAGCACC  780
GTGCGGTACT GGATGGCCTT CAACGTGGAG CGGGGCCTGCC AGTCCTACTT TGGCTGTGTG  840
CAGTGTATTA GTGGGCCCTT GGGCATGTAC CGCAACAGCC TCCTCCAGCA GTTCCTGGAG  900
GACTGGTACC ATCAGAAGTT CCTAGGCAGC AAGTGCAGCT TCGGGGATGA CCGGCACCTC  960
ACCAACCGGAG TCCTGAGCCT TGGCTACCGA ACTAAGTATA CCGCGCGCTC CAAGTGCCTC  1020
ACAGAGACCC CCACTAAGTA CCTCCGGTGG CTCAACCAGC AAACCCGCTG GAGCAAGTCT  1080
TACTTCCGGG AGTGGCTCTA CAACTCTCTG TGGTTCCATA AGCACCACCT CTGGATGACC  1140
TACGAGTCAG TGGTCACGGG TTTCTTCCCC TTCTTCCTCA TTGCCACGGT TATACAGCTT  1200
TTCTACCGGG GCCGCATCTG GAACATTCTC CTCTTCCTGC TGACGGTGCA GCTGGTGGGC  1260
ATTATCAAGG CCACCTACGC CTGCTTCCTT CGGGGCAATG CAGAGATGAT CTTCATGTCC  1320
CTCTACTCCC TCCTCTATAT GTCCAGCCTT CTGCCGGCCA AGATCTTTGC CATTGCTACC  1380
ATCAACAAAT CTGGCTGGGG CACCTCTGGC CGAAAAACCA TTGTGGTGAA CTTCATTGGC  1440
CTCATTCCTG TGTCCATCTG GGTGGCAGTT CTCCTGGAGG GGCTGGCCTA CACAGCTTAT  1500
TGCCAGGACC TGTTCAGTGA GACAGAGCTA GCCTTCCTTG TCTCTGGGGC TATACTGTAT  1560
GGCTGCTACT GGGTGGCCCT CCTCATGCTA TATCTGGCCA TCATCGCCCG GCGATGTGGG  1620
AAGAAGCCGG AGCAGTACAG CTTGGCTTTT GCTGAGGTGT GA                     1662
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MPVQLTTALR VVGTSLFALA VLGGILAAYV TGYQFIHTEK HYLSFGLYGA ILGLHLLIQS   60
LFAFLEHRRM RRAGQALKLP SPRRGSVALC IAAYQEDPDY LRKCLRSAQR ISFPDLKVVM  120
VVDGNRQEDA YMLDIFHEVL GGTEQAGFFV WRSNFHEAGE GETEASLQEG MDRVRDVVRA  180
STFSCIMQKW GGKREVMYTA FKALGDSVDY IQVCDSDTVL DPACTIEMLR VLEEDPQVGG  240
VGGDVQILNK YDSWISFLSS VRYWMAFNVE RACQSYFGCV QCISGPLGMY RNSLLQQFLE  300
DWYHQKFLGS KCSFGDDRHL TNRVLSLGYR TKYTARSKCL TETFTKYLRW LNQQTRWSKS  360
YFREWLYNSL WFHKHHLWMT YESVVTGFFP FFLIATVIQL FYRGRIWNIL LFLLTVQLVG  420
IIKATYACFL RGNABMIFMS LYSLLYMSSL LPAKIFAIAT INKSGWGTSG RKTIVVNFIG  480
LIPVSIWVAV LLEGLAYTAY CQDLFSETEL AFLVSGAILY GCYWVALLML YLAIIARRCG  540
KKPEQYSLAF AEV
```

```
1          11         21         31         41         51
|          |          |          |          |          |
CACACATACG CACGCACGAT CTCACTTCGA TCTATACACT GGAGGATTAA AACAAACAAA   60
CAAAAAAAAC ATTTCCTTCG CTCCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA  120
CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCCTAA AGCGTTTCCT CGCTTGCATT  180
CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA  240
CTTGTTGAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAA TTGGGGAAAG  300
AAATATCCAA CATGTAATAG CCCAAAACAA TCTCCTATCA ATATTGATGA AGATCTTACA  360
CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAAAC ATCATTGGAA  420
```

```
        AACACATTCA TTCATAACAC TGGGAAAACA GTGGAAATTA ATCTCACTAA TGACTACCGT    480
        GTCAGCGGAG GAGTTTCAGA AATGGTGTTT AAAGCAAGCA AGATAACTTT TCACTGGGGA    540
        AAAATGCAATA TGTCATCTGA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT    600
        GAGATGCAAA TCTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAAA    660
        GGAAAAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAATTTG    720
        GATTTCAAAG CGATTATTGA TGGAGTCGAA AGTGTTAGTC GTTTTGGGAA GCAGGCTGCT    780
        TTAGATCCAT TCATACTGTT GAACCTTCTG CCAAACTCAA CTGACAAGTA TTACATTTAC    840
        AATGGCTCAT TGACATCTCC CCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT    900
        ACAGTTAGCA TCTCTGAAAG CCAGTTGGCT GTTTTTTGTG AAGTTCTTAC AATGCAACAA    960
        TCTGGTTATG TCATGCTGAT GGACTACTTA CAAAACAATT TTCGAGACGA ACAGTACAAG    1020
        TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAAGGAAG AGATTCATGA AGCAGTTTGT    1080
        AGTTCAGAAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTTACA    1140
        TGGGAAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTGCAGT TTTGTACCAG    1200
        CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTTGA CAGATGGCTA TCAAGACTTG    1260
        GGTGCTATTC TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA    1320
        TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTCGACAT GCCTACTGAT    1380
        AATCCTGAAC TTGATCTTTT CCCTGAATTA ATTGGAACTG AAGAAATAAT CAAGGAGGAG    1440
        GAAGAGGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA    1500
        AACCAAATCA GGAAAAAGGA ACCCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG    1560
        ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTCTCTGGA    1620
        AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCAGTCAC TAAATTAGCC    1680
        ACAGAAAAAG ATATTTCCTT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG    1740
        GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAACTG TTCTTAGATC TCCACATACA    1800
        AACTTGTCGG GGACTGCAGA ATCCTTAAAT ACAGTTTCTA TAACAGAATA TGAGGAGGAG    1860
        AGTTTATTGA CCAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC    1920
        GCAACTTCTG CTATCCCATT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCCTCC    1980
        GAAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC    2040
        GAAGATTCAA CTTCATCAGG TTCAGAAGAA TCACTAAAGG ATCCTTCTAT GGAGGGAAAT    2100
        GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG    2160
        AGCTTTCTCC AGACTAATTA CACTGAGATA CGTGTTGATG AATCTGAAA GACAACCAAG    2220
        TCCTTTTCTG CAGGCCCAGT GATGTCACAG GGTCCCTCAG TTACAGATCT GGAAATGCCA    2280
        CATTATTCTA CCTTTGCCTA CTTCCCAACT GAGGTAACAC CTCATGCTTT TACCCCATCC    2340
        TCCAGACAAC AGGATTTGGT CTCCACGGTC AACGTGGTAT ACTCGGCAGAC AACCCAACCG    2400
        GTATACAAATG GTGAGACACC TCTTCAACCT TCCTACAGTA GTGAAGTCTT TCCTCTAGTC    2460
        ACCCCTTTGT TGCTTGACAA TCAGATCCTC AACACTACCC CTGCTGCTTC AAGTAGTGAT    2520
        TCGGCCTTGC ATGCTACGCC TGTATTTCCC AGTGTCGATG TGTCATTTGA ATCCATCCTG    2580
        TCTTCCTATG ATGGTGCACC TTTGCTTCCA TTTTCCTCTG CTTCCTTCAG TAGTGAATTG    2640
        TTTCGCCATC TGCATACAGT TTCTCAAATC CTTCCACAAG TTACTTCAGC TACCGAGAGT    2700
        GATAAGGTGC CCTTGCATGC TTCTCTGCCA GTGGCTGGGG GTGATTTGCT ATTAGAGCCC    2760
        AGCCTTGCTC AGTATTCTGA TGTGCTGTCC ACTACTCATG CTGCTTCAGA GACGCTGGAA    2820
        TTTGGTAGTG AATCTGGTGT TCTTTATAAA ACGCTTATGT TTTCTCAAGT TGAACCACCC    2880
        AGCAGTGATA CCATGATGCA TGCACGTTCT TCAGGGCCTG AACCTTCTTA TGCCTTGTCT    2940
        GATAATGAGG GCTCCCAACA CATCTTCACT GTTTCTTACA GTTCTGCAAT ACCTGTGCAT    3000
        GATTCTGTGG GTGTAACTTA TCAGGGTTCC TTATTTAGCG GCCCTAGCCA TATACCAATA    3060
        CCTAAGTCTT CGTTAATAAC CCCAACTGCA TCATTACTGC AGCCTACTCA TGCCCCTCTCT    3120
        GGTGATGGGG AATGGTCTGG AGCCTCTTCT GATAGTGAAT TTCTTTTACC TGACACAGAT    3180
        GGGCTGACAG CCCTTAACAT TTCTTCACCT GTTTCTGTAG CTGAATTTAC ATATACAACA    3240
        TCTGTGTTTG GTGATGATAA TAAGGCGCTT TCTAAAAGTG AAATAATATA TGGAAATGAA    3300
        ACTGAACTGC AAAATTCCTTC TTTCAATGAG ATGGTTTACC CTTCTGAAAG CACAGTCATG    3360
        CCCAACATGT ATGATAAATGT AAATAAGTTG AATGCGTCTT TACAAGAAAC CTCTGTTTCC    3420
        ATTTCTAGCA CCAAGGGCAT GTTTCCAGGG TCCCTTGCTC ATACCACCAC TAAGGTTTTT    3480
        GATCATGAGA TTAGTCAAGT TCCAGAAAAT AACTTTTCAG TTCAACCTAC ACATACTGTC    3540
        TCTCAAGCAT CTGGTGACAC TTCGCTTAAA CCTGTGCTTA GTGCAAACTC AGAGCCAGCA    3600
        TCCTCTGACC CTGCTTCTAG TGAAATGTTA TCTCCTTCAA CTCAGCTCTT ATTTTATGAG    3660
        ACCTCAGCTT CTTTTAGTAC TGAAGTATTG CTACAACCTT CCTTTCAGGC TTCTGATGTT    3720
        GACACCTGC TTAAAACTGT TCTTCCAGCT GTGCCCAGTG ATCCAAATATT GGTTGAAACC    3780
        CCCAAAGTTG ATAAAAATTAG TTCTACAATG TTGCATCTCA TTGTATCAAA TTCTGCTTCA    3840
        AGTGAAAACA TGCTGCACTC TACATCTGTA CCAGTTTTTG ATGTGTCGCC TACTTCTCAT    3900
        ATGCACTCTG CTTCACTTCA AGGTTTGGACC ATTTCCTATG CAAGTGAGAA ATATGAACCA    3960
        GTTTTGTTAA AAAGTGAAAG TTCCCACCAA GTGGTACCTT CTTTGTACAG TAATGATGAG    4020
        TTGTTCCAAA CGGCCAATTT GGAGATTAAC CAGGCCCATC CCCAAAAGG AAGGCATGTA    4090
        TTTGCTACAC CTGTTTTATC AATTGATGAA CCATTAAATA CACTAATAAA TAAGCTTATA    4140
        CATTCCGATG AAATTTTAAC CTCCACCAAA AGTTCTGTTA CTGGTAAGGT ATTTGCTGGT    4200
        ATTCCAACAG TTGCTTCTGA TACATTTGTA TCTACTGATT ATTCTGTTCC TATAGGAAAT    4260
        GGGCATGTTG CCATTACAGC TGTTTCTCCC CACAGAGATG GTTCTGTAAC CTCAACAAAG    4320
        TTGCTGTTTC CTTCTAAGGC AACTTCTGAG CTGAGTCATA GTGCCAAATC TGATGCCGGT    4380
        TTAGTGGGTG GTGGTGAAGA TGGTGACACT GATGATGATG GTGATGATGAC    4440
        AGAGGTAGTG ATGGCTTATC CATTCATAAG TGTATGTCAT GCTCATCCTA TAGAGAATCA    4500
        CAGGAAAAGG TAATGAATGA TTCAGACACC CACGAAAACA GTCTTATGGA TCAGAATAAT    4560
        CCAATCTCAT ACTCACTATC TGAGAATTCT GAAGAAGATA ATAGAGTCAC AAGTGTATCC    4620
        TCAGACAGTC AAACTGGTAT GGACAGAAGT CCTGGTAAAT CACCATCAGC AAATGGGCTA    4680
        TCCCAAAAGC ACAATGATGG AAAAGAGGAA AATGACATTC AGACTGGTAG TGCTCTGCTT    4740
        CCTCTCAGCC CTGAATCTAA AGCATGGGCA GTTCTGACAA GTGATGAAGA AAGTGGATCA    4800
        GGGCAAGGTA CCTCAGATAG CCTTAATGAG AATGAGACTT CCACAGATTT CAGTTTTGCA    4860
        GACACTAATG AAAAAGATGC TGATGGGATC CTGGCAGCAG CTGGACTCAGA AATAACTCCT    4920
        GGATTCCCAC AGTCCCCAAC ATCATCTGTT ACTAGCGAGA ACTCAGAAGT GTTCCACGTT    4980
        TCAGAGGCAG AGGCCAGTAA TAGTAGCCAT GAGTCTCGTA TTGGTCTAGC TGAGGGGTTG    5040
        GAATCCGAGA AGAAGGCAGT TATACCCCTT GTGATCGTGT CAGCCCTGAC TTTTATCTGT    5100
        CTAGTGGTTC TTGTGGGTAT TCTCATCTAC TGGAGGAAAT GCTTCCAGAC TGCACACTTT    5160
        TACTTAGAGG ACAGTACATC CCCTAGAGTT ATATCCACAC CTCCAACACC TATCTTTCCA    5220
        ATTTCAGATG ATGTCGGAGC AATTCCAATA AAGCACTTTC CAAAGCATGT TGCAGATTTA    5280
        CATGCAAGTA GTGGGTTTAC TGAAGAATTT GAGACACTGA AAGAGTTTTA CCAGGAAGTG    5340
        CAGAGCTGTA CTGTTGACTT AGGTATTACA GCAGACAGCT CCAACCACCC AGACAACAAG    5400
        CACAAGAATC GATACATAAA TATCGTTGCC TATGATCATA GCAGGGTTAA GCTAGCACAG    5460
```

(5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80)

```
        CTTGCTGAAA AGGATGGCAA ACTGACTGAT TATATCAATG CCAATTATGT TGATGGCTAC   5520
        AACAGACCAA AAGCTTATAT TGCTGCCCAA GGCCCACTGA AATCCACAGC TGAAGATTTC   5580
        TGGAGAATGA TATGGGAACA TAATGTGGAA GTTATTGTCA TGATAACAAA CCTCGTGGAG   5640
        AAAGGAAGGA GAAAATGTGA TCAGTACTGG CCTGCCGATG GGAGTGAGGA GTACGGCAAC   5700
        TTTCTGGTCA CTCAGAAGAG TGTGCAAGTG CTTGCCTATT ATACTGTGAG GAATTTTACT   5760
        CTAAGAAACA CAAAAATAAA AAAGGGCTCC CAGAAAGGAA GACCCAGTGG ACGTGTGGTC   5820
        ACACAGTATC ACTACACGCA GTGGCCTGAC ATGGGAGTAC CAGAGTACTC CCTGCCAGTG   5880
        CTGACCTTTG TGAGAAAGGC AGCCTATGCC AAGCGCCATG CAGTGGGGCC TGTTGTCGTC   5940
        CACTGCAGTG CTGGAGTTGG AAGAACAGGC ACATATATTG TGCTAGACAG TATGTTGCAG   6000
        CAGATTCAAC ACGAAGGAAC TGTCAACATA TTTGGCTTCT TAAAACACAT CCGTTCACAA   6060
        AGAAATTATT TGGTACAAAC TGAGGAGCAA TATGTCTTCA TTCATGATAC ACTGGTTGAG   6120
        GCCATACTTA GTAAAGAAAC TGAGGTGCTG GACAGTCATA TTCATGCCTA TGTTAATGCA   6180
        CTCCTCATTC CTGGACCAGC AGGCAAAACA AAGCTAGAGA AACAATTCCA GCTCCTGAGC   6240
        CAGTCAAATA TACAGCAGAG TGACTATTCT GCAGCCCTAA AGCAATGCAA CAGGGAAAAG   6300
        AATCGAACTT CTTCTATCAT CCCTGTGGAA AGATCAAGGG TTGGCATTTC ATCCCTGAGT   6360
        GGAGAAGGCA CAGACTACAT CAATGCCTCC TATATCATGG GCTATTACCA GAGCAATGAA   6420
        TTCATCATTA CCCAGCACCC TCTCCTTCAT ACCATCAAGG ATTTCTGGAG GATGATATGG   6480
        GACCATAATG CCCAACTGGT GGTTATGATT CCTGATGGCC AAAACATGGC AGAAGATGAA   6540
        TTTGTTTACT GGCCAAATAA AGATGAGCCT ATAAATTGTG AGAGCTTTAA GGTCACTCTT   6600
        ATGGCTGAAG AACACAAATG TCTATCTAAT GAGGAAAAAC TTATAATTCA GGACTTTATC   6660
        TTAGAAGCTA CACAGGATGA TTATGTACTT GAAGTGAGGC ACTTTCAGTG TCCTAAATGG   6720
        CCAAATCCAG ATAGCCCCAT TAGTAAAACT TTTGAACTTA TAAGTGTTAT AAAAGAAGAA   6780
        GCTGCCAATA GGGATGGGCC TATGATTGTT CATGATGAGC ATGAAGAGAGT GACGGCAGGA   6840
        ACTTTCTGTG CTCTGACAAC CCTTATGCAC CAACTAGAAA AAGAAAATTC CGTGGATGTT   6900
        TACCAGGTAG CCAAGATGAT CAATCTGATG AGGCCAGGAG TCTTTGCTGA CATTGAGCAG   6960
        TATCAGTTTC TCTACAAAGT GATCCTCAGC CTTGTGAGCA CAAGGCAGGA AGAGAATCCA   7020
        TCCACCTCTC TGGACAGTAA TGGTCAGCA TTGCCTGATG GAAATATAGC TGAGAGCTTA   7080
        GAGTCTTTAG TTTAACACAG AAAGGGGTGG GGGGACTCAC ATCTGAGCAT TGTTTTCCTC   7140
        TTCCTAAAAT TAGGCAGGAA AATCAGTCTA GTTCTGGTTAT CTGTTGATTT CCCATCACCT   7200
        GACAGTAACT TTCATGACAT AGGATTCTGC CGCCAAATTT ATATCATTAA CAATGTGTGC   7260
        CTTTTTGCAA GACTTGTAAT TTACTTATTA TGTTTGAACT AAAATGATTG AATTTTACAG   7320
        TATTTCTAAG AATGGAATTG TGGTATTTTT TTCTGTATTG ATTTTAACAG AAAAATTTCAA   7380
        TTTATAGAGG TTAGGAATTC CAAACTACAG AAAATGTTTG TTTTTAGTGT CAAATTTTTA   7440
        GCTGTATTTG TAGCAATTAT CAGGTTTGCT AGAAAATATAA CTTTTAATAC AGTAGCCTGT   7500
        AAATAAAACA CTCTTCCATA TGATATTCAA CATTTTACAA CTGCAGTATT CACCTAAAGT   7560
        AGAAATAATC TGTTACTTAT TGTAAATACT GCCCTAGTGT CTCCATGGAC CAAATTTATA   7620
        TTTATAATTG TAGATTTTTA TATTTTACTA CTGAGTCAAG TTTTCTAGTT CTGTGTAATT   7680
        GTTTAGTTTA ATGACGTAGT TCATTAGCTG GTCTTACTCT ACCAGTTTTC TGACATTGTA   7740
        TTGTGTTACC TAAGTCATTA ACTTTGTTTC AGCATGTAAT TTTAACTTTT GTGGAAAATA   7800
        GAAATACCTT CATTTTGAAA GAAGTTTTTA TGAGAATAAC ACCTTACCAA ACATTGTTCA   7860
        AATGGTTTTT ATCCAAGGAA TTGCAAAAAT AAATATAAAT ATTGCCATTA AAAAAAAAAA   7920
        AAAAAAAAAA AAAAAAAAAA AAAA
```

Seq ID NO: 372 Protein sequence:
Protein Accession #: built from XP_031379

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        MRILKRFLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALNQKNWG KKYPTCNSPK    60
        QSPINIDEDL TQVNVNLKKL KFQGWDKTSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV   120
        FRASKITFHW GKCNMSSDGS EHSLEGQKFP LEMQIYCFDA DRFSSPEEAV KGKGKLRALS   180
        ILFEVGTEEN LDFKAIIDGV ESVSRFGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC   240
        TDTVDNIVFK DTVSISESQL AVFCEVLTMQ QSGYVMLMDY LQNNFREQQY KFSRQVFSSY   300
        TGKEEIHEAV CSSEPENVQA DPENYTSLLV TWERPRVVYD TMIEKPAVLY QQLDGEDQTK   360
        HEPLTDGYQD LGAILNNLLP NMSYVLQIVA SDQLIVDMPT DNPKLDLFPE   420
        LIGTEEIIKE EEEGKDIEEG AIVNPGRDSA TNQIRKKBPQ ISTTTHYNRI GTKYNEAKTN   480
        RSPTRGSEFS GKGDVPNTSL NSTSQPVTKL ATEKDISLTS QTVTELPPHT VEGTSASLND   540
        GSKTVLRSPH MNLSGTAESL NTVSITEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS   600
        ENISQGYIFS SENPETITYD VLIPESARMA SEDSTSSGSE ESLKDPSMEG NVNFPSSTDI   660
        TAQPTDVGSGR ESFLQTNYTE IRVDESEKTT KSFSAGPVMS QGPSVTDLEM FHYSTFAYFP   720
        TEVTPHAFTP SSRQQDLVST VNVVYSQTTQ PVYNGETPLQ PSYSSEVFPL VTPLLLDNQI   780
        LNTTPAASSS DSALHATPVF PSVDVSFESI LSSYDGAPLL PFSSASPSSE LFRHLHTVSQ   840
        ILPQVTSATE SDKVFLHASL PVAGGDLLLE PSLAQYSDVL STTHAASETL EPGSESGVLY   900
        KTLMFSQVEP PSSDAMMHAR SSGPEPSYAL SDNEGSQHIF TVSYSSAIPV HDSVGVTYQG   960
        SLFSGPSHIP IPKSSLITPT ASLLQPTHAL SGDGEWSGAS SDSEFLLPDT DGLTALNISS  1020
        PVSVAEFTYT TSVFGDDNKA LSKSEIIYGN ETELQIPSFN EMVYPSESTV MPNMYDNVNK  1080
        LNASLQETSV SISSTKGMFP GSLAHTTTKV FDHEISQVPE NNFSVQPTHT VSQASGDTSL  1140
        KFVLSANSEP AESDPASSEM LSPSTQLLFY ETSASFSTEV LLQPSFQASD VDTLLKTVLP  1200
        AVPSDPILVE TPKVDKISST MLHLIVSNSA SSENMLHSTS VPVFDVSPTS HMHSASLQGL  1260
        TISYASEKYE PVLLKSESSH QVVPSLYSND ELFQTANLEI NQAHPPKGRH VFATPVLSID  1320
        EPLNTLINKL IHSDEILTST KSSVTGKVFA SDPTVASDTF VSTDHSVPIG NGHVAITAVS  1380
        PHRDGSVTST KLLFPSKATS ELSHSAKSDA GLVGGGEDGD TDDDGDDDDD DRGSDGLSIH  1440
        KCMSCSSYRE SQEKVMNDSD THENSLMDQN NPISYSLSEN SEEDNRVTSV SSDSQTGMDR  1500
        SPGKSPSANG LSQKHNDGKE ENDIQTGSAL LPLSPESKAW AVLTSDESSG SGQGTSDSLN  1560
        ENETSTDFSF ADTNEKDADG ILAAGDSEIT PGPPQSPTSS VTSENSEVFR VSSAEASNSS  1620
        HESRIGLAEG LESEKKAVIP LVIVSALTPI CLVVLVGILI YWRKCFQTAH FYLEDSTSPR  1680
        VISTPPTPIF PISDDVGAIP IKHFPKHVAD LHASSGFTEE FETLKEFYQE VQSCTVDLGI  1740
        TADSSNHPDN KHKNRYINIV AYDHSRVKLA QLAEKDGKLT DYINANYVDG YNRPKAYIAA  1800
        QGPLKSTAED FHRMIWEHNV EVIVMITNLV EKGRRKCDQY WPADGSEEYG NFLVTQKSVQ  1860
        VLAYYTVRNF TLRNTKIKKG SQKGRPSGRV VTQYHYTQWP DMGVPEYSLP VLTFVRKAAY  1920
        AKRHAVGPVV VHCSAGVGRT GTYIVLDSML QQIQHEGTVN IFGFLKHIRS QRNYLVQTEE  1980
        QYVFIHDTLV EAILSKETEV LDSHIHAYVN ALLIPGPAGK TKLEKQFQLL SQSNIQQSDY  2040
        SAALKQCNRE KNRTSSIIPV ERSRVGISSL SGEGTDYINA SYIMGYYQSN EFIITQHPLL  2100
        HTIKDFWRMI WDHNAQLVVM IPDGQNMAED EFVYWPNKDE PINCESFKVT LMAEEHKCLS  2160
```

```
NEEKLIIQDF ILEATQDDYV LEVRHFQCPK WPNPDSPISK TFELISVIKE EAANRDGPMI  2220
VHDERGGVTA GTFCALTTLM HQLEKENSVD VYQVAKMINL MRPGVFADIE QYQFLYKVIL  2280
SLVSTRQEEN PSTSLDSNGA ALPDGNIAES LESLV
```

5   Seq ID NO: 373 DNA sequence
    Nucleic Acid Accession #: built from NM_002851
    Coding sequence: 148-4518

```
     1         11        21        31        41        51
     |         |         |         |         |         |
    CACACATACG CACGCACGAT CTCACTTCGA TCTATACACT GGAGGATTAA AACAAACAAA   60
    CAAAAAAAAC ATTTCCTTCG CTCCCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA  120
    CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCCTAA AGCGTTTCCT CGCTTGCATT  180
    CAGCTCCTCT GTGTTTGCCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA  240
    CTTGTTGAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAAA TTGGGGAAAG  300
    AAATATCCAA CATGTAATAG CCCAAAACAA TCTCCTATCA ATATTGATGA AGATCTTACA  360
    CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGGTT GGGATAAAAC ATCATTGGAA  420
    AACACATTCA TTCATAACAC TGGGAAAACA GTGGAATTTA ATCTCACTAA TGACTACCGT  480
    GTCAGCGGAG GAGTTTCAGA AATGGTGTTT AAAGCAAGCA AGATAACTTT TCACTGGGGA  540
    AAATGCAATA TGTCATCTGA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT  600
    GAGATGCAAA TCTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAAA  660
    GGAAAAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAATTTG  720
    GATTTCAAAG CGATTATTGA TGGAGTCGAA AGTGTTAGTC GTTTTGGGAA GCAGGCTGCT  780
    TTAGATCCAT TCATACTGTT GAACCTTCTG CCAAACTCAA CTGACAAGTA TTACATTTAC  840
    AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT  900
    ACAGTTAGCA TCTCTGAAAA CCAGTTGGCT GTTTTTTGTG AAGTTCTTAC AATGCAACAA  960
    TCTGGTTATG TCATGCTGAT GGACTACTTA CAAAACAATT TTCGAGAGCA ACAGTACAAG  1020
    TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAAGGAAG AGATTCATGA AGCAGTTTGT  1080
    AGTTCAGAAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTTACA  1140
    TGGGAAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGCAGT TTTGTACCAG  1200
    CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTTGA CAGATGGCTA TCAAGACTTG  1260
    GGTGCTATTC TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA  1320
    TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTCGACAT GCCTACTGAT  1380
    AATCCTGAAC TTGATCTTTT CCCTGAATTA ATTGGAACTG AAGAAATAAT CAAGGAGGAG  1440
    GAAGAGGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGCTACA  1500
    AACCAAATCA GGAAAAAGGA ACCCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG  1560
    ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTCTCTGGA  1620
    AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCAGTCAC TAAATTAGCC  1680
    ACAGAAAAAG ATATTTCCTT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG  1740
    GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAACTG TTCTTAGATC TCCACATATG  1800
    AACTTGTCGG GGACTGCAGA ATCCTTAAAT ACAGTTTCTA TAACAGAATA TGAGGAGGAG  1860
    AGTTTATTGA CCAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC  1920
    GCAACTTCTG CTATCCCATT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTCCTCC    1980
    GAAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC  2040
    GAAGATTCAA CTTCATCAGG TTCAGAAGAA TCACTAAAGG ATCCTTCTAT GGAGGGAAAT  2100
    GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG  2160
    AGCTTTCTCC AGACTAATTA CACTGAGATA CGTGTTGATG AATCTGAGAA GACAACCAAG  2220
    TCCTTTTCTG CAGGCCCAGT GATGTCACAG GGTCCCTCAG TTACAGATCT GGAAATGCCA  2280
    CATTATTCTA CCTTTGCCTA CTTCCCAACT GAGGTAACAC CTCATGCTTT TACCCCATCC  2340
    TCCAGACAAC AGGATTGGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCAACCG  2400
    GTATACAATG CAGAGGCCAG TAATAGTAGC CATGAGTCTC GTATTGGTCT AGCTGAGGGG  2460
    TTGGAATCCG AGAAGGAAGGC AGTTATACCC CTTGTGATCG TGTCAGCCCT GACTTTTATC  2520
    TGTCCAGTGG TTCTTGTGGG TATTCTCATC TACTGGAGGA AATGCTTCCA GACTGCACAC  2580
    TTTTACTTAG AGGACAGTAC ATCCCCTAGA GTTATATCCA CACCTCCAAC ACCTATCTTT  2640
    CCAATTTCAG ATGATGTCGG AGCAATTCCA ATAAAGCACT TTCCAAAGCA TGTTGCAGAT  2700
    TTACATGCAA GTAGTGGGTT TACTGAAGAA TTTGAGACAC TGAAAGAGTT TTACCAGGAA  2760
    GTGCAGAGCT GTACTGTTGA CTTAGGTATT ACAGCAGACA GCTCCAACCA CCCAGACAAC  2820
    AAGCACAAGA ATCGATACAT AAATATCGTT GCCTATGATC ATAGCAGGGT TAAGCTAGCA  2880
    CAGCTTGCTG AAAAGGATGG CAAACTGACT GATTATATCA ATGCCAATTA TGTTGATGGC  2940
    TACAACAGAC CAAAAGCTTA TATTGCTGCC CAAGGCCCAC TGAAATCCAC AGCTGAAGAT  3000
    TTCTGGAGAA TGATATGGGA ACATAATGTG GAAGTTATTG TCATGATAAC AAACCTCGTG  3060
    GAGAAAGGAA GGAGAAAATG TGATCAGTAC TGGCCTGCCG ATGGGAGTGA GGAGTACGGG  3120
    AACTTTCTGG TCACTCAGAA GAGTGTGCAA GTGCTTGCCT ATTATACTGT GAGGAATTTT  3180
    ACTCTAAGAA ACACAAAAAT AAAAAAGGGC TCCCAGAAAG GAAGACCCAG TGGAGTGTG   3240
    GTCACACAGT ATCACTACAC GCAGTGGCCT GACATGGGAG TACCAGAGTA CTCCCTGCCA  3300
    GTGCTGACCT TTGTGAGAAA GGCAGCCTAT GCCAAGCGCC ATGCAGTGGG GCCTGTTGTC  3360
    GTCCACTGCA GTGCTGGAGT TGGAAGAACA GGCACATATA TTGTGCTAGA CAGTATGTTG  3420
    CAGCAGATTC AACACGAAGG AACTGTCAAC ATATTTGGCT TCTTAAAACA CATCCGTTCA  3480
    CAAAGAAATT ATTTGGTACA AACTGAGGAG CAATATGTCT TCATTCATGA TGCTATTTGGTT 3540
    GAGGCCATAC TTAGTAAAGA AACTGAGGTG CTGGACAGTC ATATTCATGC CTATGTTAAT  3600
    GCACTCCTCA TTCCTGGACC AGCAGGCAAA ACAAAGCTAG AGAAACAATT CCAGCTCCTG  3660
    AGCCAGTCAA ATATACAGCA GAGTGACTAT TCTGCAGCCC TAAAGCAATG CAACAGGGAA  3720
    AAGAATCGAA CTTCTTCTAT CATCCCTGTG GAAAGATCAA GGGTTGGCAT TTCATCCCTG  3780
    AGTGGAGAAG GCACAGACTA CATCAATGCC TCCTATATCA TGGGCTATTA CCAGAGCAAT  3840
    GAATTCATCA TTACCCAGCA CCCTCTCCTT CATACCATCA AGGATTTCTG GAGGATGATA  3900
    TGGGACCATA ATGCCCAACT GGTGGTTATG ATTCCTGATG GCCAAAACAT GGCAGAAGAT  3960
    GAATTTGTTT ACTGGCCAAA TAAAGATGAG CCTATAAAT GTGAGAGCTT TAAGGTCACT  4020
    CTTATGGCTG AAGAACACAA ATGTCTATCT AATGAGGAAA AACTTATAT TCAGGACTTT  4080
    ATCTTAGAAG CTACACAGGA TGATTATGTA CTTGAAGTGA GGCACTTTCA GTGTCCTAAA  4140
    TGGCCAAATC CAGATGGCCC CATTAGTAAA ACTTTTGAAC TTATAAGTGT TATAAAAGAA  4200
    GAAGCTGCCA ATAGGGATGG GCCTATGATT GTTCATGATG AGCATGGAGG AGTGACGGCA  4260
    GGAACTTTCT GTGCTCTGAC AACCCTTATG CACCAACTAG AAAAAGAAAA TTCCGTGGAT  4320
    GTTTACCAGG TAGCCAAGAT GATCAATCTG ATGAGGCCAG GAGTCTTTGC TGACATTGAG  4380
    CAGTATCAGT TTCTCTACAA AGTGATCCTC AGCCTTGTGA GCACAAGGCA GGAAGAGAAT  4440
```